# The utility of molecular genetic analysis of museum specimens in studying deep-sea fish

## Thesis

AMRA KAZIC

# The utility of molecular genetic analysis of museum specimens in studying deep-sea fish

A Thesis Presented for the Research Degree of
Doctor of Philosophy (Ph.D.)

April 2008

**Sponsoring Establishment:** NESCOT, Surrey, U.K.
**Collaborating Establishment:** The Natural History Museum, U.K.
**PhD Degree Registration:** Research School, The Open University, U.K.
**Discipline:** Life and Biomolecular Sciences (Molecular Genetics)



*This thesis is dedicated to the memory of my parents,*

*Seida Kazic (1932 - 1994)*

*Ibrahim Kazic (1929 - 2007)*

# Contents

# APPENDICES

# Acknowledgements

This voyage took many years, and I will not be able to express my gratitude to everyone who directly and indirectly contributed to this thesis and its completion.

First of all, I would like to thank my main supervisor and Director of my Ph.D. study Dr John B.W. Hammond (NESCOT) for giving me the chance to work on this project, guiding me through laboratory work and development of the thesis towards submission, proof-reading of the thesis, and also for his kind support and encouragement throughout. I am also very grateful to my Ph.D. supervisor Dr Nigel R. Merrett (now retired) from the Natural History Museum (NHM) for advice and information on the general taxonomic and sample preservation aspects of the study, and for kindly proof-reading sections of this thesis. Dr David A. Johnston (NHM) helped me with sequence analysis and sequence validation. I would like to thank Clive Moncrieff (a science statistician/Head of Biometric Group at the NHM, Department of Zoology, Section for Biometry and Molecular Computing Research) for helping me with statistics and statistical analysis of my data. Also, I would like to thank Professor Philip S. Rainbow (NHM) for critical reading of the manuscript of my thesis, helpful comments and corrections, and for his other support and help as a Head of Zoology Department in the Natural History Museum, London.

I would like to thank Oliver Crimmen, Patrick Campbell and James MacLaine – fish curators from the Natural History Museum, London for helping me to understand preservation processes in fish and aspects of general sample preservation.

I am grateful to Dr Alex Rogers from the Oceanography Centre, Southampton University (now in the Institute of Zoology, Zoological Society of London) for providing me with ethanol preserved tissue samples of *Coryphaenoides,* helpful advice and information on deep-sea organisms and their genetics, and giving me the opportunity to visit his Laboratory for two weeks in 1999 in order to test and validate some DNA extraction protocols and PCR primers.

This work was sponsored by the North East Surrey College of Technology (NESCOT), Ewell, Epsom, Surrey U.K., including my 3-yr studentship and Ph.D. registration fee at the Open University. I am grateful to my collaborating establishment, the Natural History Museum (NHM), London, U.K. for providing the museum specimens for this study, financial support (NHM Special Fund) for *Scomber* study, library and software resources, laboratory facilities for validation of results, and other help and support. I gratefully acknowledge the time and support given by my employer, the Institute for Genetic Engineering and Biotechnology (INGEB), Sarajevo, Bosnia and Herzegovina.

I would like to thank all my friends and colleges for support and assistance. So many people helped me pull through the academic and emotional challenges, but a special thank you to my dear friend Barbara Anderson (Academic Registrar at Gresham College, London) for giving me enormous moral support, listening to me for hours, and for her professional advice and help. Also, I would like to thank to Dr Jeni Stewart from the NHM for her kind advice and assistance.

I wish to express my warmest thanks and much appreciation to my sister for supporting me in many ways during all these years.

STATEMENT 1

This thesis is the result of my own investigations, except where otherwise stated. Other sources are acknowledged by giving explicit references. A bibliography is appended.

STATEMENT 2

I hereby give consent for my thesis, if accepted, to be made available for photocopying and for inter library loan, and for the title and summary to be made available to outside organisations.

# List of publications and presentations related to this research

**Kazic A., Hammond J.B.W., Merrett N.R. and Crimmen O. (2007):** An investigation of the utility of molecular genetic analysis by using formalin preserved animal tissue. *5th ISABS Conference in Forensic Genetics and Molecular Antropology*, Split, Croatia, 3-7 September 2007; Book of Abstracts: p. 113
**(POSTER)**

**Kazic A., Hammond J.B.W., Merrett N.R. and Crimmen O. (2007):** DNA sequences from museum specimens of fish with unstudied genomes. *XII European Congress of Ichthyology*, Cavtat (Dubrovnik), Croatia, 9-13 September 2007; Book of Abstracts: p. 33
**(ORAL PRESENTATION)**

**Kazic A., Hammond J.B.W., Johnston D.A., Merrett N.R. and Crimmen O. (2005):** Museum collections and molecular-genetic information: two molecular studies on museum fish specimens. *3rd Annual London Evolutionary Research Network Conference (L.E.R.N.)* held at the Natural History Museum, London, U.K, 16 September 2005.
**(POSTER)**

**Kazic A., Hammond J.B.W., Johnston D.A., Merrett N.R. and Crimmen O. (2005):** Museum collections and molecular-genetic information. *SPNHC 20th Annual Meeting and Workshops 2005* held at the Natural History Museum, London, U.K, 12-18 June 2005
**(POSTER)**

**Kazic A., Hammond J.B.W., Johnston D.A., Merrett N.R. and Crimmen O. (2005):** Molecular data on preserved fish specimens from the collection of the Natural History Museum, London. *2005 PORCUPINE Annual General Meeting* (PORCUPINE Marine Natural History Society): *Collections, Collectors, Collecting* held at the Natural History Museum, London, U.K, 18-20 March 2005.
**(ORAL PRESENTATION)**

**Kazic A., Hammond J.B.W., Johnston D.A., Merrett N.R. and Crimmen O. (2004):** Use of museum fish specimens for molecular investigations. *2nd LERN (London Evolutionary Research Network) Conference* held at the Zoological Institute, London, U.K, *17th Sep 2004.*
**(ORAL PRESENTATION)**

**Kazic A., Hammond J.B.W., Johnston D.A. and Merrett N.R. (2003):** Development of reliable, reproducible methods for the molecular analysis and sequencing of formalin-fixed fish specimens. *The Second Annual Student Poster Display by PhD Students at the NHM,* London, U.K.
**(POSTER)**

**Kazic A. and Hammond J.B.W. (2000):** The first molecular data for *Nezumia* (Grenadier) museum specimens. *International Conference: Aquatic Habitats as Ecological Islands. A Meeting of the British Ecological Society Aquatic Group*, University of Plymouth, U.K.
**(POSTER)**

**Kazic A. and Hammond J.B.W. (2000):** Methodological approaches for analyzing formalin-preserved museum specimens of fish (genus *Nezumia*) using RAPD and mtDNA. *44. International Annual Conference of Ecological Genetics Group*, Edge Hill, Ormskirk, U.K.
**(ORAL PRESENTATION/ DISCUSSION TOPIC)**

**Hammond J.B.W., Kazic A. and Spanswick G. (1998):** Use of AGS Gold *Taq* Polymerase for RAPD analysis of museum specimens. *Hybaid Catalogue 1998/99*, pp. 13-15.

# List of Abbreviations and Acronyms:

| | |
|---|---|
| **aDNA** | ancient DNA |
| **arDNA** | archival DNA |
| **bp** | base pair |
| **COI** | the mitochondrial cytochrome *c* oxidase 1 gene |
| **cyt *b*** | the mitochondrial cytochrome *b* gene |
| **DMSO** | dimethylsulfoxide |
| **DNA** | deoxyribonucleic acid |
| **dNTPs** | deoxyribonucleotide triphosphates |
| **DPCs** | DNA-protein crosslinks |
| **EDTA** | ethylenediaminetetraacetic acid |
| **EFEP** | ethanol-fixed, ethanol-preserved |
| **FFEP** | formalin-fixed, ethanol-preserved |
| **FFFP** | formalin-fixed, formalin-preserved |
| **HPLC** | high-performance liquid chromatography |
| **IMS** | industrial methylated spirits |
| **ITS** | internal transcribed spacer |
| **kb** | kilobases |
| **LINE** | long interspersed repetitive elements |
| **MAAP** | multiple arbitrary amplicon profiling techniques |
| **Mb** | megabases |
| **MOPS** | 3-(N-morpholino) propanesulfonic acid ($C_7H_{15}NO_4S$) |
| **mtDNA** | mitochondrial DNA |
| **N.a** | *Nezumia aequalis* |
| **NCBI** | The National Center for Biotechnology Information |
| **nDNA** | nuclear DNA |
| **NHM** | The Natural History Museum, London, U.K. |
| **N.m** | *Nezumia micronychodon* |
| **ORF** | Open Reading Frame |
| **PCI** | phenol : chloroform : isoamyl alcohol extraction |
| **PCR** | Polymerase Chain Reaction |
| **PE** | Perkin Elmer |
| **RAPD** | Randomly Amplified Polymorphic DNA |

| | |
|---|---|
| **RFLP** | Restriction Fragment Length Polymorphism |
| **RNA** | ribonucleic acid |
| **SCAR** | Sequence Characterised Amplified Region |
| **SDS** | sodium dodecyl sulphate |
| **SINE** | short interspersed repetitive elements |
| **SNP** | single nucleotide polymorphism |
| **STS** | Sequence Tagged Site |
| **TE** | Tris-HCl EDTA buffer |
| **UV** | ultraviolet light |
| **12S** | 12S mitochondrial DNA |
| **16S** | 16S mitochondrial DNA |

# List of Tables

# List of Figures

# Support material on CD

***Support material to sections 3.1. – 3.3.***

**Tables S1 - S9:** Summarised PCR results of experiments presented in Figures S1 - S4

**Figures S1:** Agarose gel images showing RAPD-PCR experiments

**Figures S2:** Agarose gel images showing mitochondrial PCRs and PCRs with designed *Nezumia*-specific primers

**Figures S3:** Agarose gel images showing RAPD-PCRs generated by RAPD primer 44 on arDNA extracts produced by protocols (B) and (D)

**Figures S4:** Agarose gel images showing RAPD-PCR experiments performed with RAPD primer 29 on arDNA extracts produced by protocol (C) – problems with a PCR inhibition

**Figures S5 – S10:** Agarose gel images showing RAPD-PCRs performed with different RAPD primers on different fish individuals of formalin-fixed, Steedman's-preserved specimens of *Nezumia aequalis* and *N. micronychodon* (with different arDNA extracts) – Reproducibility of RAPD-PCR profiles (bands)

***Support material to section 3.2. (BLAST search results)***

**Table S10:** BLASTN search results with the 458 bp consensus DNA sequence of *Nezumia* against the GenBank (NCBI) nucleic acid database

**Table S11:** Translated BLAST search results (blastx and tblastx) with the 458 bp consensus DNA sequence of *Nezumia* against the GenBank (NCBI) database

**Table S12:** BLASTP search results with the 152 amino acids long protein sequence against the GenBank (NCBI) peptide database (458 bp long consensus DNA sequence of *Nezumia* was translated into the protein sequence by the use of the "Sequencher" software)

**Table S13:** BLASTN search results with the 358 bp cloned DNA sequence of *Nezumia aequalis* against the GenBank (NCBI) nucleic acid database

**Table S14:** Translated BLAST search results (blastx and tblastx) with the 358 bp cloned DNA sequence of *Nezumia aequalis* against the GenBank (NCBI) database

# Abstract

The damaging effect of formalin on DNA and the inhibition of PCR are serious problems in molecular studies. The aims of the project were to investigate the possibility of using formalin-fixed, Steedman's preserved museum specimens in molecular investigations, especially organisms with unstudied genomes. A number of DNA extraction protocols and different pre-washing/drying regimes were tested. These gave different levels of success, but a guanidinium-based protocol developed in this study gave the best results. RAPD-PCR methodology was employed to test its applicability on preserved specimens, and it was used as a test of the efficiency of DNA extraction/amplifications and for developing species-specific PCR primers. Attempts to amplify mitochondrial DNA sequences with the six mitochondrial genes were mostly unsuccessful. Sporadic amplifications were obtained with primers of 16S and COIII genes.

This study provided the first molecular data on deep-sea fish (*Nezumia aequalis* and *N. micronychodon*) exclusively using formalin-fixed, Steedman's preserved museum specimens. Two genomic sequences of these fishes were determined and submitted to the GenBank database under accession numbers AY826774 – AY826792. Three specific primer sets (RAPD-derived) for *Nezumia aequalis* and *N. micronychodon* were designed to amplify PCR product sizes 300 bp – 350 bp. This study has demonstrated that an appropriate strategy and molecular approach could lead to the successful use of museum and other formalin-fixed archival collections even on organisms with unstudied genomes.

Supplementary evidence, related to the method of preservation, the usage of particular DNA extraction protocol and PCR marker system, was obtained from ten differently preserved mackerel (*Scomber scombrus*) specimens.

This study confirmed that the DNA extracted from preserved specimens possesses unique characteristics that make molecular investigations very difficult. Because of this, it is proposed that DNA extracted from preserved specimens should be referred to as "***archival DNA (arDNA)***".

# Chapter 1. GENERAL INTRODUCTION

Molecular studies open up the possibility to study rare and extinct species, or those that come from inaccessible or difficult-to-sample localities, sub-fossils and preserved specimens from museum and other archival collections. In most cases, molecular techniques have high resolving power and give accurate and reliable results even in the cases of a very limited number of samples or amount of sampled material (Wandeler *et al.* 2007). Also, it is possible to revise previously obtained results (primarily species identification and relationships amongst taxa) carried out by traditional morphological methods and techniques, which is a great advantage for resolving problems in "difficult" groups or taxa of organisms.

In this molecular genetic study, the possibility of analysing formalin-fixed, Steedman-preserved museum specimens will be discussed, and the first molecular genetic data from the deep-sea fishes *Nezumia aequalis* (Gunther, 1878) and *N. micronychodon* Iwamoto, 1970, obtained exclusively from the collection of the Natural History Museum in London, presented. The data produced could be used as a basis for further studies in developing a field where preserved archival specimens could be extensively used in molecular investigations for different biological applications. Retrieved genetic information from *Nezumia* (which was used as the test material) could be further used in studies of these and related taxa from different geographic areas. There is no significant published work in which formalin-fixed samples have been used for study because of the associated degradation of DNA and, as yet, methodologies for DNA retrieval from preserved specimens to obtain reliable molecular information are undeveloped and unvalidated (Tang 2006; Gilbert *et al.* 2007b; Skage and Schander 2007; Wandeler *et al.* 2007).

Greater use of such specimens could add a valuable temporal and spatial dimension to biological studies (Boyle *et al.* 2004; Zardus *et al.* 2006). Species that are, for various reasons, difficult to sample because of low population density and inaccessibility can be studied, since specimens accumulated from a specific location over a period of years can be analysed and a large enough body of data assembled for significant conclusions to be drawn (Schander and Halanych 2003; Austin and Melville 2006; Wandeler *et al.* 2007). Some

collections may include specimens taken 100 or more years ago or species that do not inhabit particular areas any longer, including endangered or even extinct species (Goldstein and DeSalle 2003). Analysis of these compared with more recent (or remaining) specimens may lead to conclusions concerning long term population shifts, changes in gene flow and other time and space related aspects of population and ecological genetics which cannot be made in any other way. Exploring the possibility of the molecular genetic characterisation of museum type specimens (primary: holotypes, lectotypes, neotypes, syntypes; or paratypes, paralectotypes, topotypes, chirotypes, hypotypes, homeotypes, etc.) is of significant scientific interest, and it would add valuable data from existing museum collections which then could be used broadly by researchers around the world (Schander and Willassen 2005; Wandeler *et al.* 2007).

It is worthwhile exploring the possibility and viability of obtaining data by molecular analysis from museum and other collections that contain samples preserved in different ways (formalin-fixed, in ethanol, in DMSO-NaCl, dried, etc.). Information collected by this type of research could be applicable not only in biological science, but also in forensics, medicine, agriculture and so on. This kind of research is time-consuming and labour-intensive, but the discovery of effective and reliable techniques could contribute to a better understanding of population shifts, evolution and speciation (Wandeler *et al.* 2007). As more research is done, we will become more aware of the problems and how to overcome them (Tang 2006).

## *1.1. Molecular techniques in the development of taxonomy and systematics*

Over time, definitions of taxonomy, systematics, species, populations and speciation have become more "open" and flexible. Various authors present different definitions of these terms (compare: Abercrombie *et al.* 1973; Maclean 1987; Lawrence 1995; Martin and Hine 2000; Baquero 2005; Rundle and Nosil 2005), but more flexibility is noticeable in the definitions presented over the last decade (e.g. Mayr 1996; Hillis *et al.* 1996; Baum 1998; Martin and Hine 2000; Mallet and Willmott 2003; Fitzhugh 2005). Lately, new terms have been introduced or suggested to be introduced, such as a taxospecies, biochemical taxonomy, DNA taxonomy, "reverse taxonomy" (Markmann and Tautz 2005), phylogenetic taxonomy with its nomenclature (e.g. the PhyloCode; Cantino and de Queiroz 2004), or even novel biological disciplines such as "molecular systematics", "archaeological systematics", molecular barcoding, etc.

Generally, the most difficult to define is the difference between systematics and taxonomy, and also the terms species, subspecies and population which have been problematic since the 1970s, as well as at the beginning of the 21st Century. Actually, systematics and taxonomy overlap so much that it is very difficult to draw a clear line between them. It is possible to notice that some authors use these terms almost as synonyms (e.g. Maclean 1987; Lawrence 1995). Basically, *taxonomy* is the science of naming and classifying organisms (Mallet and Willmott 2003) and *systematics* is the biological discipline that describes and explains diversity in the biological world (Hillis *et al.* 1996).

The question: "What is a species and what is a subspecies?" will, in my opinion, still stay open for long time, if it could ever be answered with an exact definition. According to Hey *et al.* (2003), Sites and Marshall (2004) and many other authors (e.g. Baum 1998; Benton 2000; Sereno 2005; Fitzhugh 2005), the main problem in defining the term "species" is a conflict between two ideas – species as "categories" that are created by the biologists who study them by using particular methods of analyses, and species as real observable entities in nature. The majority of researchers accept the term species as a real natural entity and as a fundamental unit in taxonomy and in biological studies. The recent initiative for the global barcode project, to identify and classify nearly all animal life by using short DNA sequences of cytochrome *c* oxidase I (COI; Hebert *et al.* 2003) and 16S ribosomal DNA (16S; Blaxter 2004), is probably the best reflection of the scientific community in this regard (see also the website of the Consortium for the Barcode of Life (CBOL); http://barcoding.si.edu/). After much scientific debate and discussion in the last three years, researchers generally supports DNA barcoding as one of the elements in *integrative taxonomy* which uses a large number of characters, including DNA and many other types of data, to delimit, discover and identify natural species and taxa at all levels (e.g. Lipscomb *et al.* 2003; Mallet and Willmott 2003; Proudlove and Wood 2003; Bond 2004; Lee 2004; De Ley *et al.* 2005; DeSalle *et al.* 2005; Page *et al.* 2005; Will *et al.* 2005; Dasmahapatra and Mallet 2006). Molecular data using "DNA barcode" gene fragments have been effective in discovering cryptic taxa and the recognition of species of many different groups of organisms (e.g. Janzen *et al.* (2005) on Lepidoptera; Monaghan *et al.* (2005) on tropical beetles; Page *et al.* (2005) on cryptic Australian freshwater shrimps *Caridina*; Saunders (2005) on red macroalgae; Ivanova *et al.* (2007) on fishes; Moura *et al.* (2008) on marine hydroids, and so on). Some problems are expected for application of DNA barcoding to all animal groups (Moritz and Cicero 2004), but there are already suggestions for its general improvements (e.g. developing nuclear barcodes as a supplement to the mtDNA-based barcodes; Dasmahapatra and Mallet 2006). The "DNA barcode" project is challenging and scientifically interesting from many aspects.

It might bring much more information than just its application for species identification and biodiversity assessment.

The barcode project intends to include preserved museum specimens (Hajibabaei *et al.* 2005; Janzen *et al.* 2005; De Ley *et al.* 2005; Schander and Willassen 2005), but the analysis of archival collections is currently challenging due to the problem of DNA degradation (Hajibabaei *et al.* 2005; Tang 2006). However, previous research and methods that were used and developed to retrieve DNA from archival preserved specimens (e.g. Su *et al.* 1999; Junqueira *et al.* 2002; Rohland *et al.* 2004) have been useful for barcode purposes, as well as PCR cocktails that combine polymerases with repair enzymes (Skage and Schander 2007). Recently, the use of mini-barcodes (~100 bp) was suggested for museum specimens, have already proved to be successful for some oven-dried (age 2-21 years) and ethanol-preserved (age 1-14 years) insect specimens/species (Hajibabaei *et al.* 2006).

Reliable species identification and "good" taxonomy are of crucial scientific importance for conducting different types of biological studies (Beattle and Ollver 1994; Brower 1995; Campbell 1995; Payne and Sorenson 2002). Correct identification is problematic for non-taxonomists (especially for molecular biologists), as well as for taxonomists if difficult taxa are in question (Schander and Willassen 2005) and especially if taxonomists are not experts in particular groups of organisms (e.g. for non-taxonomists, and even for ichthyologists, macrourid deep-sea fishes are a difficult group for morphological identification; Iwamoto *et al.* 1999). It can be of great benefit to isolate molecular markers and retrieve species-specific DNA sequences for such groups of organisms.

### 1.1.1. The usefulness of molecular-genetic methods for species delimitation and solving other biological problems

Contemporary biological science strives to answer a number of questions concerning biodiversity and the identification of new species, population structuring, genetics, speciation, migration patterns of different groups of organisms, systematic and phylogenetic relationships, evolution, and so on. Most researchers are of the opinion that multiple sources of data (morphology, genetics, behaviour, ecology) help best to describe the diversity of life; that is it is best first to collect morphological data, then molecular tools can provide a valuable *a posteriori* classification of differentiation at other biological levels, providing data on the correspondence between genotypic and phenotypic divergence (Lipscomb *et al.*

2003; Mallet and Willmott 2003; Bond 2004; Lee 2004; Page *et al.* 2005). Molecular methods undoubtedly provide very powerful tools in identifying taxa, assessing biodiversity, and elucidating evolutionary and ecological relationships (Schander and Willassen 2005; Wandeler *et al.* 2007). There are studies where:

- morphological and molecular data are in concordance (McCafferty *et al.* 2002),
- discrepancy between these two strands of data are evident (Wiens and Penkrot 2002),
- traditional morphological approaches could not help in resolving particular biological problem(s) because of insufficient resolving power of applied methodologies and analysis, but molecular approaches were successful (Finnerty and Block 1995; Rodriguez-Grana *et al.* 2004),
- neither molecular or morphological analysis could resolve a particular problem in full (Cano *et al.* 2005).

Different molecular markers reveal information from different parts of the genome and they might have a different "resolving power", because of different organisation and different rates of evolution (e.g. on average, genes of mtDNA are evolving about four times faster than most of nuclear coding regions; from: Wiens and Penkrot 2002). Each study requires a specific approach and an appropriate strategy in order to solve adequately a particular problem (i.e. to answer a specific question).

## *1.2. Mitochondrial and nuclear DNA*

Some molecular studies prefer to apply investigations on mitochondrial and/or nuclear DNA – sometimes it is a personal preference of a researcher, but often it is related to the type of study (which biological question/s need to be answered) and/or the nature of biological material used in a particular study. For example fresh/frozen, preserved material, forensic samples, sub-fossils, i.e. whether tissue samples contain well preserved high-molecular-weight DNA of good yield, or degraded and damaged DNA with a poor yield.

The mitochondrial and nuclear genomes possess some specific characteristics that might influence their selection for a particular study as being more appropriate and/or informative, but a combination of mitochondrial and nuclear markers is recommended for most studies in order to obtain comprehensive and informative results (Medina *et al.* 1999; Ross *et al.* 1999; Hoarau *et al.* 2004; Gaines *et al.* 2005; Azeredo-Espin and Lessinger 2006). The results

derived from mitochondrial and nuclear genomes might be in concordance (Peijnenburg *et al.* 2005), but not necessarily so and often they are not (Medina *et al.* 1999; Sanetra and Crozier 2003). This is because of differences in the structural organisation of organellar and nuclear genomes, differences in their modes of inheritance, differences in rates of evolution of mitochondrial and nuclear DNA, as well as differences in the functions and rates of evolution of particular regions of genomes and loci (Lynch *et al.* 2006).

### 1.2.1. Mitochondrial genome

Mitochondrial DNA (mtDNA) is a small, closed, double-stranded, circular molecule located within the mitochondria of eukaryotes, usually present with 1000-10000 copies in a cell (Machugh *et al.* 2000; Brown 2002). Because of the characteristic that mtDNA occurs in a higher copy number in comparison with nuclear DNA, mitochondrial DNA is preferably used in molecular research with specimens that contain degraded and therefore a small amount of DNA. It is more likely that mtDNA will be retrieved from samples of preserved specimens (France and Kocher 1996; Boyle *et al.* 2004; Zardus *et al.* 2006), archaeological biological material and sub-fossils (Hanni *et al.* 1990; Cooper *et al.* 1997; Sorenson *et al.* 1999a; Sica *et al.* 2002), *post mortem,* histopathology and forensic samples (Budowle *et al.* 2003; Alonso *et al.* 2005), and other "difficult" samples, such as: faeces and hair (Taberlet and Bouvet 1994; Lucchini *et al.* 2002), fish scales and otoliths (Durand *et al.* 1999).

Mitochondrial DNA varies extraordinarily in size, gene content and genome organisation between different groups of organisms. Animal mtDNAs are typically small molecules - 14 to 20 kb (Brown 1985; Boore 1999; Lynch *et al.* 2006), but much larger animal mitochondrial genomes (from 20 kb and up to 42 kb) have been found in some molluscs (Snyder *et al.* 1987; Gjetvaj *et al.* 1992; Rigaa *et al.* 1995), nematodes (Beck and Hyman 1988), insects (Boyce *et al.* 1989), and other invertebrates. Vertebrates generally show a smaller and less variable size range of the mitochondrial genome (Lynch *et al.* 2006). It falls in the range of 16 to 18 kb (Meyer 1993; Lynch *et al.* 2006), but there are vertebrates with significant size variations and mitochondrial genomes larger than 20 kb, especially amongst the lower vertebrates (for instance, in snakes: Kumazawa and Nishida (1995); in lizards and frogs: Macey *et al.* (1997); in fish: Gach and Brown (1997), Nesbo *et al.* (1998)). Fungi and plants generally have much larger and hugely different mtDNA sizes than animals. Fungi mitochondrial genome sizes range from ~ 19 to 95 kb, whereas plant mitochondrial genome size falls in the range from ~ 180 to 600 kb (from: Lynch *et al.* 2006).

The traditional view of the animal mitochondrial genome has been as haploid and non-recombining (Hayashi *et al.* 1985), strictly maternally or clonally inherited, conserved in genome size, gene content and gene order (Rand 1993, 1994), and rapidly evolving (Meyer 1993; Hellberg 2006; Lynch *et al.* 2006). However, recent comparative studies on animal mtDNA provided new insights into mitochondrial genome which started to change/modify some of these "traditional" views on mtDNA and animal mitochondrial genomes (Macey *et al.* 1997; Burger *et al.* 2003a,b; Endo *et al.* 2005).

Generally, animal mitochondrial genomes are highly streamlined with a compact and efficient organisation (Meyer 1993; Lynch *et al.* 2006). With few exceptions (e.g., the brachiopod *Lingula anatina*; Endo *et al.* 2005), all animal mitochondrial genomes contain the same 37 genes (13 encoding for proteins, 22 for transfer RNAs and two for ribosomal RNAs) and one major non-coding region termed as the displacement-loop (D-loop) which contains a control region (Cantatore and Saccone 1987; Meyer 1993; Boore 1999; Dowton 2004; Lynch *et al.* 2006). Genes are separated by very few nucleotides (and overlap in places) and, generally, there are no introns (Moritz *et al.* 1987; Meyer 1993; Gray *et al.* 1998; Burger *et al.* 2003a,b). It was suggested that the compactness of the mitochondrial genome is partly responsible for the stability of gene order in vertebrates, since the lack of introns and large intergenic spacers makes the mitochondrial genome less likely to rearrange (from: Meyer 1993). However, comparative analysis carried out in the last decade revealed that size variations of mtDNA (size variant heteroplasmy) might affect gene order; that is, it might cause gene rearrangements (Macey *et al.* 1997; Miya and Nishida 1999; Inoue *et al.* 2003; Mueller and Boore 2005).

Animal mtDNA, including human, is thought to be strictly maternally inherited (Giles *et al.* 1980; Birky 1995) and that its inheritance is clonal because of active degradation of paternal mitochondria during fertilization (Vaughn *et al.* 1980), or its "outreplication" shortly thereafter (Meland *et al.* 1991). However, a growing literature now indicates transmission of the paternal mtDNA into the egg and its survival in the adult organism (usually in very small amounts) in a number of species (e.g., Kondo *et al.* (1990) in *Drosophila*; Gyllensten *et al.* (1991) in mice; Zouros *et al.* (1992) in the marine mussel; Magoulas and Zouros (1993) in anchovy; Kvist *et al.* (2003) in birds), including humans (Schwartz and Vissing 2002), although nuclear-encoded proteins exist which target sperm mitochondria in oocytes for their destruction (Cummins *et al.* 1998; Shitara *et al.* 1998, 2000). This suggests that genetic recombination may occur in animal mtDNA in variety of taxa (Thyagarajan *et al.*

1996; Lunt and Hyman 1997; Kajander *et al.* 2001; Ladoukakis and Zouros 2001; Hoarau *et al.* 2002; Rokas *et al.* 2003; Kraytsberg *et al.* 2004; Endo *et al.* 2005; Tsaousis *et al.* 2005; Guo *et al.* 2006). Besides proved biparental mitochondrial inheritance in many animal organisms, there is growing evidence that genetic recombination (between different haplotypes) in mtDNA within and between populations, subspecies and species might occur (Piganeau *et al.* 2004; Tsaousis *et al.* 2005).

From an evolutionary aspect, animal mitochondrial and nuclear genomes have evolved in different directions (Lynch *et al.* 2006). The evolution of the nuclear genome has included expansion in size, number of introns, and increased lengths of intergenic regions, whereas animal mitochondrial genomes have evolved in opposite directions scaling down in size and reducing the number and size of non-coding regions (Gray *et al.* 1998; de Grey 2005; Lynch *et al.* 2006). Different genomic architecture and evolution have produced mitochondrial and nuclear DNA-specific genetic systems that differ in a number ways (Lynch *et al.* 2006).

In many organisms, the mtDNA accumulates mutations more rapidly than do single-copy nuclear genes (Brown *et al.* 1979; Meyer 1993; Gissi *et al.* 2000; Wan *et al.* 2004; Hellberg 2006; Lynch *et al.* 2006). In other words, mtDNA provides markers with greater variability and sensitivity to drift and other factors than nDNA, and is therefore more likely to show differences between populations/species, i.e. indicate a population/species subdivision (Nyakaana *et al.* 2002). This makes the mitochondrial genome attractive for systematic (Shedlock *et al.* 1992; Banks *et al.* 1993; Moran *et al.* 1994; Munasinghe *et al.* 2003; Harris *et al.* 2005) and population genetic studies (Avise *et al.* 1986; Reeb and Avise 1990; Zwanenburg 1992; Geller *et al.* 1993; Hauser *et al.* 1998; Hurwood *et al.* 2003; Aboim *et al.* 2005), as well as for investigating historically contingent questions and phylogenetic relationships (Kreitman and Wayne 1994; Scoles *et al.* 1998; Helfenbein *et al.* 2004; Simon *et al.* 1994, 2006). However, selective sweeps (Maruyama and Birky 1991), stochastic lineage extinction (Avise *et al.* 1984), and founder events (DeSalle and Templeton 1988; Gamache *et al.* 2003) may reduce mitochondrial DNA variability but have little effect on nuclear variability (Medina *et al.* 1999; Hellberg 2006).

One of the proposed reasons for higher rates of nucleotide substitutions in mtDNA vs. nDNA is the much smaller effective population size of mitochondrial genes than for nuclear genes (Birky *et al.* 1989; Sanetra and Crozier 2003). Recently, it was suggested that three factors may promote elevated DNA mutation rates in animal mitochondria (from: Lynch *et al.* 2006):

1. Mitochondria generate free oxygen radicals producing an internal environment with an exceptionally high mutagenic potential (Balaban *et al.* 2005);
2. In contrast to nuclear DNA, mitochondrial DNA is continuously replicated within non-dividing cells, and the base-misincorporation rate (before proof-reading) is ~ $10^3 - 10^4$ times higher than in the nuclear genome (Johnson and Johnson 2001);
3. Mitochondrial nucleotide-excision repair may have been entirely lost (Mason and Lightowlers 2003).

It is assumed that mtDNA might be more vulnerable than nDNA (*in vivo* and *in vitro*) to hydrolytic and oxidative damage because of the lack of wound histone proteins which presumably absorb some of the surrounding damage (Lindahl 1993b; Mecocci *et al.* 1994; Poinar 2002). It was proven that mtDNA is a critical cellular target for reactive oxygen species (ROS); that is, 3-fold more damage to the mitochondrial genome has been observed in comparison to nuclear DNA if treated with hydrogen peroxide ($H_2O_2$) for 15 and 60 minutes (Yakes and van Houten 1997). They also found that, following 60-min treatment, damage to the nuclear DNA was completely repaired within 1.5 hours, whereas no DNA repair in the mitochondrion was observed. This might be of the significance in mtDNA preservation and the integrity of the mtDNA molecule after the organism dies (natural decay), i.e. there may be a differential preservation of mtDNA and nDNA (see Berger *et al.* 2001). Furthermore, these might be important and related to mtDNA preservation in archival collections; that is, the integrity, accessibility and extractability of mtDNA from chemically-preserved organisms deposited in the museums and other archival collections.

#### *1.2.1.1. Mitochondrial mutational and damage hotspots*

Some recent publications on ancient mtDNA (Gilbert 2003a,b; Gilbert *et al.* 2005b) suggest possible correlations between the distribution of mtDNA *post-mortem* damage ("damage hotspots"), on the one hand, and the regions (sites) of modern mtDNA known to have high *in vivo* mutation rates, on the other hand. Further investigations with regard to the correlation between "*in vivo* mutation and *post-mortem* damage hotspots" might provide interesting information not only on ancient DNA, but also on the nature and distribution of damage in DNA extracted from preserved and other "difficult" samples.

What causes particular regions of mtDNA to be more susceptible to *in vivo* mutations and which cellular/genetic mechanisms regulate this? Are these presumably "highly mutational" regions of mtDNA also equally more vulnerable and susceptible to DNA damage (base modifications, strand breakages, cross-links) induced by death of organisms (decay) and other factors (chemicals used in preservation of tissue/organisms, exposure to: UV, heat, acidic or alkali pH, etc.)? These are extremely interesting questions and relate to the fundamentals of genetics and biological science. Hypotheses about the existence of *in vivo* mutation hotspots and a hotspot of gene order rearrangements in mtDNA is already supported by some evidence and models (e.g. Galtier *et al.* (2006) on mutation hotspots in mammalian mitochondrial coding regions; San Mauro *et al.* (2006) on a hotspot of gene order rearrangement in the vertebrate mitochondrial genome). The advents of emulsion-based clonal amplification (emPCR) and sequencing-by-syntheses technology have also provided new insights into aDNA miscoding lesions and their distribution, supporting the hypothesis suggesting the existence of damage hotspots in aDNA (Gilbert *et al.* 2007a). Banerjee and Brown (2004) have reported on non-random artifactual sequence changes in a 181 bp segment of the mitochondrial *atp*A gene (adenine to guanine change in particular) induced by heat treatment (wheat DNA was heated at 95°C for 2-21 days). Brandstatter *et al.* (2005) has reported the existence of DNA sequencing error hotspots in mtDNA in clinical, anthropological and forensic samples which are mainly related to laboratory-specific factors (the sort of automated sequencer and sequencing chemistry employed, and other laboratory-specific factors, such as post-PCR treatments).

Inevitably, there is an immediate set of questions to be asked considering a broader issue related to *in vivo* DNA, ancient and otherwise degraded and damaged DNA: "Do these *in vivo* "mutational and gene rearrangements hotspots" found in mtDNA also exist in the nDNA? If these hotspots exist in the nDNA, is the "hotspot hypotheses" established for mtDNA also partly, or equally, applicable to nDNA – i.e. are the regions of nDNA prone to mutation also more vulnerable to damage induced by a death of organism and/or other factors (chemical preservation, exposure to the UV, high temperature, radiation, etc.)? Which genetic mechanisms and factors could be involved?" Some of these issues related to nuclear aDNA have already started to be explored (Binladen *et al.* 2006).

### *1.2.1.2. Approaches in studying mtDNA*

Although mitochondrial genomics - mitogenomics (sequencing and comparative analysis of complete mitochondrial genome sequences of different organisms) and genomics in general (sequencing and comparative analysis of whole genomes) are nowadays technically possible and are important approaches for solving complex biological problems (Curole and Kocher 1999; Cooper *et al.* 2001; Lan *et al.* 2003; Boore *et al.* 2005; Wilson *et al.* 2005), the study of single mitochondrial genes (i.e. fragments of genes) is still a broadly applied approach in contemporary molecular research (including DNA barcoding). This "traditional" molecular approach is valuable and applicable for addressing many different biological questions (see: Johns and Avise 1998; Sunnucks 2000, and previous pages).

Mitochondrial genes can easily be amplified by the polymerase chain reaction (PCR) in almost every organism by using universal mitochondrial primers for targeting highly conserved regions of mtDNA (Kocher *et al.* 1989; DeSalle *et al.* 1993b; Sorenson *et al.* 1999b), or mitochondrial primers specifically developed for a particular group of organisms and/or taxa. The extraction of mtDNA from fresh, frozen and short-term ethanol preserved tissue/specimens is now a routine procedure for many organisms, but the extraction of mtDNA might be difficult if extracted from biological material other than fresh/frozen tissue (France and Kocher 1996; Junqueira *et al.* 2002; Chakraborty *et al.* 2006). Even frozen material might cause a reduced yield of mtDNA (Gjetvaj *et al.* 1992; Berger *et al.* 2001). Berger *et al.* (2001) reported preferential harm to mtDNA in comparison to nDNA.

#### 1.2.1.2.1. Formalin-fixed specimens in studying mtDNA

Formalin fixation of specimens/tissue samples usually causes considerable problems in studying mtDNA (Li *et al.* 2000; Collins *et al.* 2002; Diaz-Viloria *et al.* 2005; Boore *et al.* 2005; Chakraborty *et al.* 2006). There are published papers on successful extractions and investigations of mtDNA using formalin-fixed (but ethanol-preserved) tissue samples (for instance: France and Kocher (1996) on crustaceans with 16S and CO I genes; Shedlock *et al.* (1997) on fish with 16S and cyt b genes; Chase *et al.* (1998a,b) and Boyle *et al.* (2004) on molluscs with 16S, COI and cyt b), but there is also information on unsuccessful attempts to extract mtDNA from formalin-fixed specimens stored for prolonged periods (or even for a few weeks) in formalin and/or ethanol including Finnerty and Block (1995),

Diaz-Viloria *et al.* (2005) and Chakraborty *et al.* (2006) on fish. Because mtDNA occurs in many copies per cell, successful mtDNA extractions and molecular investigations might be expected using preserved specimens, but the issue seems to be much more complex and not completely understood. It might be related to a differential preservation of nDNA and mtDNA in fluid preserved specimens, difficulties in releasing mtDNA from cross-linked mitochondria and protein-DNA complexes, as well as faster degradation of mtDNA in general (Berger *et al.* 2001). Possible differential preservation of nuclear and organellar DNA has already been suggested for bioarchaeological samples and plant material (nuclear and chloroplast DNA in 2000-year-old plant remains and dry plant leaves up to 289 days old; Banerjee and Brown 2002).

Difficulties in extracting mtDNA from formalin-fixed tissue samples might also be related to the functions of mitochondria and mitochondrial involvement in cell death. Disruptions in mitochondrial functions and death of cells create a highly "reactive" environment with high levels of free radicals in the mitochondria (peroxide radicals [$^{\cdot}O_2$], hydrogen peroxide [$H_2O_2$], and hydroxy radicals [$^{\cdot}OH$]) that can cause oxidation of DNA ("oxidative attacks on mtDNA") and lead to endogenous DNA damage, such as: DNA-protein cross-links, strand breaks, possible chemical alterations in DNA sequence, and so on (Cadet *et al.* 1997; Johansen *et al.* 2005). Any disruption in the transport of electrons, ions and chemicals in mitochondria might result in a release of proteins in the cytoplasm and in the inter-membrane space of mitochondria, as well as in cytoplasmic acidification (Desagher and Martinou 2000; Newmeyer and Ferguson-Miller 2003). By death of cells, the concentration of proteins, including enzymes inside mitochondria and in its outer membrane, is increased (Newmeyer and Ferguson-Miller 2003; Huang *et al.* 2004). These increased amounts of proteins in and around the mitochondria (especially in its membranes) might create grounds for the increase of DNA-protein and protein-protein complexes during chemical interactions with formalin (used for fixation/preservation of tissue/specimens) which potentially, can make the release of mtDNA (mtDNA extractability) very difficult. The formation of cross-links induced by formalin (*in vivo* and *in vitro*) has been experimentally proved and demonstrated by many researchers, including the specificity of proteins that become crosslinked to DNA (e.g., Solomon and Varshavsky 1985; Orlando *et al.* 1997; Kaufman *et al.* 2000; Barker *et al.* 2005b; Dai *et al.* 2005; Yamashita 2007; Bogenhagen *et al.* 2008).

For example, Kaufman *et al.* (2000) experimentally demonstrated *in organello* formaldehyde cross-linking of proteins to mtDNA by using 1% formaldehyde at 4°C for 16 hours. They also demonstrated that shorter exposures of mitochondria to formaldehyde

result in a lower level of protein-mtDNA cross-linking. Furthermore, Kaufman *et al.* (2000) proved the possibility of reversing some protein-mtDNA cross-links (e.g. Abf2 protein-mtDNA cross-links) by treating them with 1% sarkosyl in Tris buffer at 95°C for 30 minutes, but not by treating them with 1% SDS or 1% SDS plus 30 seconds of heating treatment. This supports the suggestion that the exposure of mitochondria to formalin influences the degree of cross-linking, i.e. the reversibility of the process. Their experiments might also indicate the importance of every step in mtDNA extraction procedures for archival preserved specimens (choice of chemicals and their concentrations, the order of their use, etc.) in order to succeed in releasing mtDNA from the "cross-linked mitochondria" and crosslink complexes in formalin-treated tissues.

### 1.2.2. Nuclear genome

Genome size varies enormously among organisms ranging from 1.2 x $10^7$ bp (the yeast *Saccharomyces cerevisiae*) to >6 x $10^{11}$ bp (*Amoeba dubia*) (from NCBI genome database). In teleost fish, the smallest genomes are in "fugu" and pufferfish (385-400 Mb) and among the largest ones are salmonid fishes (2600-3000 Mb). Nuclear DNA (nDNA) contains coding and non-coding DNA regions. It is estimated that only 1%-3% of nucleotides in the mammalian genome regulates or codes for essential proteins (Clay *et al.* 2003). Whereas coding regions can exhibit high degrees of sequence homology in the same gene between distantly related organisms, non-coding sequences are often more variable (from: Park and Moran 1995). Some parts of the genome are more conserved than others and they have different rates of change through time and under specific conditions (see review of Volff (2005) for genome evolution and biodiversity in teleost fish).

It is possible to distinguish repetitive and non-repetitive DNA (also called a single-copy nuclear DNA; scnDNA). Non-repetitive nDNA is present only once in the haploid genome, whereas repetitive DNA can be repeated from a few to a few hundred times in the nuclear genome (Liu and Cordes 2004) and is usually non-coding, but not exclusively (for instance, ribosomal genes). Repetitive DNA can be divided into two classes: tandemly repeated sequences and interspersed repeats. One form of repetitive DNA is called satellite DNA with repeat units ranging in length from one to a few hundred nucleotides (base pairs), repeated many times in tandem arrays (Schlotterer 2004). Based on the sizes of the repeat units, it is possible to distinguish minisatellites and microsatellites. Minisatellites contain repeat units in length from ten to a hundred nucleotides (Jeffreys *et al.* 1985), whereas

microsatellites (called also "simple sequence repeats"; SSR) contain smaller repeat units - usually one to four nucleotides long, and are variable as well (Tautz 1989; Kimmel *et al.* 1998; Harr and Schlotterer 2000; Schlotterer 2000). Satellite DNA can be located anywhere in the genome although new evidence suggests that the genomic distribution of microsatellites is not completely random as was originally thought (Li *et al.* 2004). In terms of function, satellite DNA is considered to be non-coding (Britten and Kohne 1968; Stephan and Cho 1994), but many recent studies have found that a large number of microsatellites are located in transcribed regions of genomes (i.e. in open reading frames - ORFs), including protein-coding genes and expressed sequence tags (Edwards *et al.* 1998; Tompa 2003; Li *et al.* 2002, 2004; Woodhead *et al.* 2005).

Transposable elements (TE) are another type of repetitive DNA. They are repetitive mobile sequences that are dispersed throughout the genome, but, repeated copies are not tandemly repeated (Singer 1982; Smit 1999; Prak and Kazazian 2000; Comeron 2001). There are two classes of transposable elements: DNA transposons and retrotransposons (Prak and Kazazian 2000). Transposable elements can also be classified according to their degree of mechanic self-sufficiency: *autonomous*, such as long interspersed nuclear elements (LINE-1 or L1 retrotransposons), and *non-autonomous*, such as short interspersed nuclear elements (SINEs) (Prak and Kazazian 2000). Autonomous TEs encode essentially all of the machinery that they require to move because they contain open reading frames (ORFs), whereas non-autonomous TEs are entirely dependent on other transposable elements for their mobility because they lack an ORF (Prak and Kazazian 2000). Autonomous TEs seem to insert preferentially into A+T-rich regions, whereas non-autonomous TEs accumulate in G+C-rich (gene-rich) DNA regions (Smit 1999; Prak and Kazazian 2000).

Repetitive nDNA is also the ribosomal DNA multigene family (18S, 5S and 28S ribosomal genes in animals) together with its spacers: two internal transcribed spacers (ITS-1 and ITS-2), an external transcribed spacer (ETS), and a nontranscribed spacer (NTS). This cluster can be tandemly repeated in several hundred copies in a eukaryote nuclear genome (Long and Dawid 1980; Hillis and Dixon 1991). From a functional aspect, Long and Dawid (1980) consider this gene repetition as "dosage repetition" because the cell requires a large amount of the particular product which a single copy could not produce in the appropriate time. As a molecular marker, this cluster is potentially a suitable candidate for different types of investigations because different regions of the rDNA repeat unit evolve at different rates (Hillis and Dixon 1991; Medina *et al.* 1999; Ran *et al.* 2001; Ji *et al.* 2003).

### *1.2.2.1. The possibility of studying nuclear genomes by the use of specimens with degraded DNA*

Nuclear single/low copy genes are rarely used for molecular investigations of specimens with degraded and damaged DNA (sub-fossils, preserved specimens, non-invasive samples, etc.) because of difficulties of their recovery from such samples (Wandeler *et al.* 2007). However, ribosomal genes with their spacers can be potentially useful in molecular work with "difficult" samples and unstudied genomes because of the higher possibility of their recovery - occurrence in many copies in the nuclear genome, and the possibility to apply universal PCR primers for particular ribosomal DNA (rDNA) regions (Rocco *et al.* 2005). Some researchers have already investigated their application to preserved and other difficult specimens (Li *et al.* 2000; Boyle *et al.* 2004; Roeder *et al.* 2004; Bhadury *et al.* 2005, 2006a,b; Mtambo *et al.* 2006; Karaiskou *et al.* 2007; Skage and Schander 2007). The amplifications were of variable success according to their published results. For preserved specimens, it depended on the type of preservation and length of storage, species (group of organisms), region of the ribosomal cluster used for investigations and size of amplification product, on the one hand, and DNA extraction protocol and pre-extraction treatment of preserved tissue, on the other hand. Overall, their results suggest that it is worthwhile to explore the applicability of this genome region and associated markers to archival specimens and other difficult samples.

There is potential for using microsatellites for investigating preserved and other difficult specimens if repetitive regions are known and if PCR primers are already developed for particular taxa (Austin and Melville 2006; Wan *et al.* 2006; Watts *et al.* 2007). However, problems, such as incorrect genotyping and "allelic dropouts", are expected (or even an inability to use them for some specimens) due to DNA damage and DNA degradation - see Taberlet *et al.* (1999); Morin *et al.* (2001); Siwoski *et al.* (2002); Boyle *et al.* (2004); Buchan *et al.* (2005); Hoffman and Amos (2005); Dewoody *et al.* (2006).

Transposon display (Waugh *et al.* 1997) and "anchored polymerase chain reaction" (Ayyadevara *et al.* 2000), i.e. sequence-specific amplification polymorphism (S-SAP), are novel techniques and marker systems associated with repetitive DNA transposon elements which it might be possible to use on degraded DNA if conserved regions of transposons are known for investigated taxa (see Patton *et al.* 2001; Porceddu *et al.* 2002; Behura *et al.* 2004), or for unstudied genomes if other techniques and methodologies (such as RAPD) are

initially employed for detecting transposable elements (see Abe *et al.* 2000). Their use has still not been assessed on preserved samples.

The application of multiple arbitrary amplicon profiling (MAAP) techniques, such as RAPD (random amplified polymorphic DNA; Williams *et al.* 1990), is possible to use on preserved specimens, but problems related to the correctness of profiling (non-reproducibility and inconsistency of RAPD results) are expected (see Eckerman and Walsh 1997; Carvalho and Vieira 2000; Oliveira *et al.* 2002; Siwoski *et al.* 2002; Boyle *et al.* 2004).

## *1.3. Molecular techniques and molecular markers*

Molecular techniques made it possible to carry out scientific investigations and obtain genetic information on different organisms and biological materials that are not accessible by "traditional" methods and techniques (Wandeler *et al.* 2007). This also enables revision and re-validation of existing conclusions previously made by morphological and less sensitive molecular analyses that may lead to correction in the classification (Payne and Sorenson 2002; Malhotra and Thrope 2004; Rodriguez-Grana *et al.* 2004) and/or in the evolutionary pattern of some groups of organisms and taxa (DeSalle 1994; Huynen *et al.* 2003).

The molecular revolution has come in waves. The first wave started with the widespread use of ***protein*** analysis, the next started with the development of ***recombinant DNA technology*** and centred on the use of restriction fragment length polymorphism (RFLPs) and sequencing of DNAs cloned in bacteria. The present wave of technology is based on the ***amplification of DNA sequences*** enzymatically via the Polymerase Chain Reaction (PCR) and sequencing of whole genomes.

The purpose of all developed molecular methods is to detect variation, differences and similarities in DNA sequences between individuals, populations, species and higher taxa. In general, they are based on one of three different classes of markers: protein variants, DNA sequence polymorphism and DNA repeat variation (Schlotterer 2004).

### 1.3.1. Restriction Fragment Length Polymorphism (RFLP)

The discovery and isolation of restriction endonucleases in the 1960s set the grounds for developing direct, DNA-based markers which can survey DNA variation itself, rather than rely on indirect, enzyme-based markers. In the late 1970s, the application of restriction enzymes was first applied to the analysis of mitochondrial DNA (Avise 1978; Avise *et al.* 1979). This was the beginning of the development of a new class of genetic marker - restriction fragment length polymorphism (RFLP) and restriction site analysis, which became widely accepted in fundamental molecular research and applied science (Nei and Tajima 1981; Berg and Ferris 1984; Wirgin *et al.* 1997; Branco *et al.* 2000; Sanjuan and Comesana 2002; Fajardo *et al.* 2006). Restriction fragment length polymorphism is due to a mutation that results in a change in the pattern of restriction fragments generated when a DNA molecule is treated with a restriction enzyme (Hale *et al.* 1995).

Development of the referent fragments for RFLP analysis and its application on degraded and damaged DNA might be difficult and not fully reliable because of fragmented DNA and possible artificial base (nucleotide) modifications induced by post-mortem effects and/or some chemicals (see Karaiskou *et al.* (2007) for application of RFLP to formalin preserved fish eggs). Furthermore, the application of the RFLP on samples with degraded DNA might be difficult because of difficulties in finding restriction sites in small DNA fragments and in visualising (detecting) very small RFLP fragments on the gel.

### 1.3.2. Polymerase Chain Reaction (PCR)

The invention of the polymerase chain reaction (PCR) in the 1980s (Saiki *et al.* 1985, 1988; Mullis and Faloona 1987) was the main breakthrough for developing PCR-based methodologies and DNA molecular markers. The polymerase chain reaction makes it possible to analyse the tiny amounts of DNA that are preserved in some sub-fossils and bioarchaeological material (e.g. Paabo *et al.* (1988) - 7000-year old brain; Brown *et al.* (1994) - 700-2000-years old wheat seeds; Loreille *et al.* (2001) - 20,000-130,000 years old bear bones), as well as other specimens with degraded and damaged DNA, such as:

- museum and other archival collections (e.g. Ellegren (1991) on dry museum birds; Shedlock *et al.* (1997) on formalin-fixed, ethanol-preserved museum fish specimens up to 85 years old; Townson *et al.* (1999) on dry entomology museum collections

ranging in age from 15 to 93 years; Moran and Baker (2002) - fish scales stored in mounting cards; Boyle *et al.* (2004) - formalin-fixed specimens of bivalves up to a few decades old),

- herbarium fungal and plant material (Savolainen *et al.* 1995; Ristaino 1998; Ristaino *et al.* 2000; Cozzolino *et al.* 2006);
- processed food (e.g. Mackie *et al.* (1995) for identification of canned fish species);
- forensic, post-mortem and histopathological medical specimens (e.g., Ludes *et al.* 1993; Romero *et al.* 1997; Jovanovic *et. al.* 2003; Edson *et al.* 2004);
- excreta of animals and humans (e.g., Gerloff *et al.* 1995; Vu *et al.* 1999; Valiere and Taberlet 2000).

The usefulness of PCR technology is also evident in genetic studies of very small organisms and their individual identification (e.g. harpacticoid copepods; Schizas *et al.* 1997), or individual identification of insects, marine larva and fish eggs (Silberman and Walsh 1992; Bilodeau *et al.* 1999; van Bortel *et al.* 2000; Kirby and Reid 2001; Perez *et al.* 2005; Aranishi 2006), as well as in forensic-agricultural applications (e.g. Congiu *et al.* (2000) for identifying strawberry varieties).

The simplest definition of the PCR is an *in vitro* method for producing large amounts of a specific, targeted DNA fragment of defined length from small amounts of DNA template (from: White 1996). The DNA piece is usually specified by two short DNA fragments (primers) which are chemically synthesized to match the base sequences flanking the sequence of interest. However, some PCRs, such as RAPD-PCRs, require only one PCR primer for generating a set of DNA fragments from anonymous regions of the genome (Williams *et al.* 1990). Or, multiplex PCRs require more than two PCR primers for generating two or more specific DNA fragments (PCR products) (see Gilbert *et al.* 2007c). PCR technology has become irreplaceable in many areas of fundamental and applied biology.

#### *1.3.2.1. Problems with PCR technology, especially in conjunction with degraded and damaged DNA-PCR templates*

PCR technology is very sensitive and some problems might be expected in this regard (Palumbi 1996a, Wilson 1997), especially if applied to impure and degraded DNA samples such as ancient DNA (Hansen *et al.* 2001; Hofreiter *et al.* 2001a,b) or DNA extracted from preserved specimens (Vachot and Monnerot 1996; Faulkner and Leigh 1998; Akbari *et al.*

2005). The fidelity of PCR might be problematic in some investigations. The possible causes are of different nature (age and post mortem changes, chemical preservation), but one of them is the fidelity of the thermostable DNA polymerase itself. According to Goodman and Tippin (2000) and Goodman (2002), the Y-family bypass DNA polymerases insert non-template bases across damaged templates (translesional synthesis), exhibit low replication fidelity and lack 3` exonuclease activity ("proofreading"). The A-family of DNA polymerases, to which *Taq* DNA polymerase belongs (commonly used for PCR), exhibits translesional synthesis with damaged DNA templates (Smith *et al.* 1998; Duarte *et al.* 2000; Patel *et al.* 2001). Novel thermostable Y-family polymerases (Dpo4-like enzymes), able to bypass DNA lesions and preferentially incorporate the correct nucleotides opposite each lesion, are developed specifically for application to damaged DNA (McDonald *et al.* 2006). At the moment, the recommendation of these authors is to blend a conventional *Taq* polymerase with a newly developed Dpo4-like thermostable Y-family polymerase in order to improve the PCR fidelity with difficult DNA templates.

The error rate of thermostable DNA polymerase is usually between $2.1 \times 10^{-4}$ to $1.6 \times 10^{-6}$ per nucleotide per cycle (Keohavong and Thilly 1989; Eckert and Kunkel 1990, 1991). The highest fidelity was reported for the thermostable DNA polymerase derived from *Pyrococcus furiosus* (*Pfu*) with an error rate $1.6 \times 10^{-6}$ per base (Lundberg *et al.* 1991), $2 \times 10^{-6}$ (Flaman *et al.* 1994), or $6.5 \times 10^{-7}$ mutations per base pair duplication (Andre *et al.* 1997). The error rate of the most commonly used *Taq* polymerase (derived from *Thermus aquaticus*) is a little higher than with *Pfu* and it is usually between $\sim 2 \times 10^{-4}$ to $< 1 \times 10^{-5}$ per base duplication per cycle (Keohavong and Thilly 1989; Eckert and Kunkel 1990, 1991). Each DNA polymerase possesses specific characteristics that could be useful for particular molecular investigations, but could be problematic in other applications. For example, high fidelity polymerases ("proofreading" DNA polymerases, such as *Vent* and *Pfu*) are good for increasing the fidelity of PCR product and correct incorporation of the bases during PCR, but they might halt PCR amplification if the PCR-DNA template is damaged and/or modified (Greagg *et al.* 1999).

The choice of which polymerase to use in PCR is important, not only for the fidelity of enzymatic DNA amplifications, but also for the amplification efficiency (Abu Al-Soud and Radstrom 1998; Sharma *et al.* 2002; Arezi *et al.* 2003; Quach *et al.* 2004; Bhadury *et al.* 2005). There are many factors and *in vitro* conditions of DNA synthesis that increase, or decrease, PCR yield and base errors in amplicons (DNA-PCR copies). For example, lower concentrations of $MgCl_2$ and dNTPs, short synthesis time with a higher annealing

temperature should increase the PCR fidelity (Wilson 1997). Jones *et al.* (1997) suggest that $Mg^{2+}$ concentration and annealing temperature are particularly important because high $Mg^{2+}$ concentration and low annealing temperature might produce "non-specific" (spurious) bands.

These are some of the experimental approaches for controlling the fidelity of PCR amplifications, but the quantity, purity and quality of DNA template are of significant importance for the accuracy of the PCR amplifications, that is, PCR fidelity (see Golenberg *et al.* 1996; Quach *et al.* 2004). Akbari *et al.* (2005) demonstrated that low DNA template input in PCR can generate false mutations by incorrect base incorporation, mainly guanine to adenine transitions. They argue that this phenomenon is not only related and exclusively caused by microheterogeneity of the used sample material (DNA samples that contain a mixture of differently damaged/altered DNA sequences) because the same artifacts were seen after amplification of a homogenous sample material that was highly diluted ("good" DNA, but of a very low concentration). Akbari *et al.* (2005) performed tests on formalin-fixed, paraffin-embedded gastric tumour tissue as microheterogenic samples, and diluted plasmid as a homogenous sample material. Mahony *et al.* (2000) and Akbari *et al.* (2005) demonstrated that a low amount of DNA template (i.e. a low number of DNA molecules available at the beginning of the PCR amplification) decreases reproducibility and the fidelity of PCR amplifications regardless of the quality of a PCR template. Impure and damaged DNA-PCR templates could only further decrease the fidelity of PCR amplifications (Taberlet *et al.* 1999; Quach *et al.* 2004).

There are reports about formalin-induced infidelity in PCR-amplified DNA fragments/sequences (e.g. De Giorgi *et al.* 1994; Williams *et al.* 1999; Quach *et al.* 2004) and artificial mutations (misincorporations) induced by post-mortem effects and age in ancient DNA (Hansen *et al.* 2001; Hofreiter *et al.* 2001a,b; Gilbert 2003a; Threadgold and Brown 2003; Pusch and Bachmann 2004; Mitchell *et al.* 2005; Gilbert *et al.* 2005a, 2007a; Stiller *et al.* 2006). For specimens from natural history collections, estimates so far range from 17% to 21% of sequences that show one or more errors, but for formalin-preserved specimens, the frequency of sequence alterations is presumed to be even higher (Wandeler *et al.* 2007). The number of sequence alterations induced by formalin and by *Taq* polymerase infidelity was reported by Williams *et al.* (1999) as being approximately one mutation artifact per 500 bases. Quach *et al.* (2004) recorded 3 to 4 times more mutations with DNA from formalin-preserved clinical specimens than with DNA extracted from a fresh tissue. Hofreiter *et al.* (2001b) suggest that DNA sequence errors induced by damage

in ancient DNA (aDNA) are unlikely to be more frequent than 0.1%, even under the unlikely scenario where amplification starts from a single template molecule. Improvements have been made in reducing PCR errors and increasing the PCR fidelity (Di Bernardo *et al.* 2002; Kaur and Makrigiorgos 2003; Pruvost *et al.* 2005), but PCR-based molecular investigations with samples that contain degraded and damaged DNA are still problematic and challenging (Cooper 2006; Tang 2006; Gilbert *et al.* 2007b; Skage and Schander 2007).

The power and sensitivity of the PCR technology, such as in detection and amplification of targeted nucleotide sequences from only a few (or even single) copy of targeted DNA in a DNA-PCR template (Cooper 2006), are advantages of the technology, but, at the same time, the source of problems and complications (Yang and Watt 2005). One of PCR phenomena related to aDNA (and, probably, to other "kinds" of low-yield and degraded DNA) is the formation of chimeric amplification molecules (Paabo *et al.* 1990). Contaminating DNA could be mixed with ancient DNA, which by "jumping PCR" (Paabo *et al.* 1990) might produce chimeric amplification sequences (DeSalle *et al.* 1993a; Gilbert 2003a). Chimeric molecules are formed when damaged ancient DNA templates are incompletely copied during a PCR cycle, allowing the incomplete extension product itself to act as a primer in the next round of PCR (Wayne *et al.* 1999); that is, when DNA polymerase jump from one short template to another to form longer but inaccurate copies of the original templates (Rogers *et al.* 2000). Generally, damaged and degraded DNA causes much more problems during PCR than a high-molecular-weight DNA of a good yield. Degraded DNA often causes erroneous insertion of bases (nucleotides), inaccurate amplification of repetitive motifs, and other artifacts of a different nature (Wayne *et al.* 1999).

Extraneous contaminant DNA can come from clothes, reagents, pipettors, plastic-ware, DNA from previous experiments or the atmosphere in genetic laboratories (Wilson 1997; Hebsgaard *et al.* 2005), as well as from bacteria, viruses, fungi or parasites that an investigated sample might contain (Paabo 1993; Prendini *et al.* 2002). Contaminant DNA and cross contamination could generate numerous copies of irrelevant DNA and results will be worthless, or, sometimes, can lead to erroneous conclusions (see Walden and Robertson 1997; Audic and Beraud-Colomb 1997; Austin *et al.* 1997; Hassanin 2002). The problem increases if genetically uncharacterised species are the subject of molecular investigation. Contaminating sequences might not be so relevant in cases where taxon-specific PCR primers are applied (primers that only match sequence of the species under investigation), but extremely complicated if universal primers for random priming (e.g. RAPD primers), or evolutionary conserved sets of specific PCR primers are applied.

PCR product carryover, which occurs when PCR products from one reaction contaminate those of another (Wilson 1997), might be serious problem in PCR experiments with any DNA, but with degraded and damaged DNA in particular (Hebsgaard *et al.* 2005; Willerslev and Cooper 2005; Wandeler *et al.* 2007). This often happens if all precautions against contamination are not respected during experimental work. In molecular work with ancient DNA, bioarchaeological and forensic samples, high standards are already set up in order to safeguard the authenticity of results, including having separate and dedicated areas for: DNA extractions, PCR set up experiments, and a room for analyses of PCR products (Brown *et al.* 1998; Cooper and Poinar 2000; Poinar 2003). With preserved specimens, this has just started to become a practice in laboratory set-ups (see Bhadury *et al.* 2006a; Wandeler *et al.* 2007).

#### 1.3.2.1.1. PCR inhibition

Problems with PCR inhibitors are likely to occur in PCR amplifications that use DNA extracted from archival collections (O`Leary *et al.* 1994; Dorris 1999; Schander and Halanych 2003; Bucklin and Allen 2004; Diaz-Viloria *et al.* 2005; Mulligan 2005; Tang 2006; Karaiskou *et al.* 2007; Skage and Schander 2007). PCR inhibition is manifested as the absence of PCR amplification (false negative), or as PCR of poor efficiency producing an extremely low quantity of the PCR product - on the gel visualised as a very faint band (Wilson 1997). PCR might be susceptible to PCR inhibition by compounds present naturally in the particular biological samples (e.g. haemoglobin and urea; Makowski *et al.* 1997), by chemicals used in fixation and preservation of organisms (Diaz-Viloria *et al.* 2005; Mulligan 2005), chemicals and compounds absorbed from the environment, and even from chemicals applied in a DNA extraction procedure (Wilson 1997). For instance, SDS, EDTA, phenol and other chemicals used in DNA extraction procedures are known as PCR inhibitors. According to the data of Qiagen (2002), strong PCR inhibitors are:

- **Phenol;** significant inhibitory effects on PCR at a final concentration of 0.2%. At 0.5% of phenol, PCR amplifications were completely inhibited and no PCR product was detectable in their research.
- **SDS (sodium dodecyl sulphate);** the yield of PCR product was decreased dramatically at a concentration of 0.005% (w/v), and it caused complete PCR inhibition at a final concentration of 0.01%. In the research of Jackson *et al.* (1990),

the negative effects of SDS were reported both on DNA extraction and PCR amplifications in molecular investigation of formalin-preserved clinical samples.

- **Sodium acetate (NaAc)** decreased the yield of PCR products at a final concentration of 5 mM, while 15 mM concentration of NaAc completely inhibited PCR.
- **EDTA (ethylenediaminetetraacetic acid)** showed an inhibitory effect on PCR and reduced the yield of PCR products at >0.5 mM, and at a final concentration of 1 mM EDTA - PCR products were not visible.

Traditionally, phenol-chloroform protocols include all, or almost all, of the above-mentioned chemicals. McNevin *et al.* (2005) reported the inhibition of PCR caused by the lysis buffer components, SDS and NaCl, when not sufficiently removed from DNA extracts. Weissensteiner and Lanchbury (1996) provided compelling evidence that as little as 10 mM NaCl inhibits the PCR of GC-rich amplicons.

Organisms that lived in a particular area might absorb specific chemicals during their lifetime (Kress *et al.* 1998; Mormede and Davies 2001; Sole *et al.* 2001). These substances might co-purify with extracted DNA and have significant effects on the success of PCRs. The experiments with bio-archaeological material are probably the best demonstration of the possible sample-surrounding effect on PCR inhibition. For example, Latham (2003) had difficulties with PCR experiments using skeletal DNA (bone remains from 19$^{th}$ and early 20$^{th}$ century) – DNA was degraded with evident post-mortem DNA damage, but PCR inhibition was also caused with substances from soil known as being PCR inhibitors (such as humic acid; Tuross 1994). Latham (2003) classified PCR inhibitors into two groups: (1) non-diffusible and (2) diffusible inhibitors. According to Latham (2003):

(1) "*Non-diffusible* inhibitors are chemical modifiers of the DNA molecule, including strand damage and molecules attached to the DNA. These modifications occur as the DNA begins to degrade after the individual's death."

(2) "*Diffusible* inhibitors are contaminating molecules not attached to DNA that leach into bone during degradation and co-purify with the skeletal DNA. These contaminating molecules come from a variety of sources and may vary among samples. The inhibitory substances leach into the bone during diagenesis and may not be removed with conventional methods used to extract skeletal DNA. Iron, tannins, and humic acid are known PCR inhibitors present in organic plant material

and soil. The presence of these substances in the immediate environment of the degrading bone may lead to their incorporation into the bone."

In order to detect the causes of PCR inhibition, Latham (2003) used bacteriophage lambda DNA as a detector of diffusible PCR inhibitors. Latham (2003) mixed lambda DNA with a small amount of extracted bone DNA, and then performed PCR amplification with lambda bacteriophage specific primers. If diffusible inhibitors are present in the bone DNA, they inhibit amplification of the lambda DNA. If that was the case, Latham (2003) applied additional purification steps in order to eliminate PCR diffusible inhibitors from a DNA sample. If a skeletal DNA sample failed to yield PCR product after the confirmed absence of diffusible PCR inhibitors (confirmed by lambda PCR-DNA experiments), Latham (2003) would presume that failure to amplify was due to non-diffusible inhibitors (DNA damage) and/or insufficient amount of DNA for PCR application.

The above explains the importance of distinguishing diffusible (contaminating molecules/compounds that co-purified with DNA) and non-diffusible (actual damage/alterations of a DNA molecule/ DNA sequence) PCR inhibitors. The described example is related to skeletal DNA, but investigating the presence of diffusible and non-diffusible inhibitors in DNA samples extracted from preserved and other difficult specimens is also crucial, although the sources of diffusible inhibitors and the nature of DNA damage (non-diffusible inhibitors) might be different (Wilson 1997; Wandeler *et al.* 2007). Mixing "good" and degraded DNA (i.e. "spiking" DNA) is often applied in molecular work with ancient and forensic DNA in order to investigate the quality of DNA extracts, i.e. the presence of diffusible PCR inhibitors (Tuross 1994; Arroyo-Pardo *et al.* 2002; Gilbert 2003a; Pusch and Bachmann 2004; Serre *et al.* 2004; von Wurmb-Schwark *et al.* 2004; Willerslev *et al.* 2004).

If chemicals that cause PCR inhibition are known to be present, the problem usually can be solved by applying procedures for their removal during DNA extractions and/or as additional purification steps after extraction (Wilson 1997). However, DNA extracts of preserved archival (Tang 2006; Wandeler *et al.* 2007) and bioarchaeological specimens (Kalmar *et al.* 2000; Arroyo-Pardo *et al.* 2002) contain mostly unknown coextracted chemicals, chemical mixtures and other substances that might act as PCR inhibitors because of variations in preservation methods, the compounds present in particular animals/plants, or variations between burial sites. Valasek and Repa (2005) suggest that the use of alternative DNA polymerases (e.g., *Tfl, Pwo, Tth,* etc.) might help in solving this problem, because

different thermostable DNA polymerases are resistant to particular inhibitors.

### *1.3.2.2. Real-time PCR technology (RT-PCR)*

In recent years, real-time PCR (Higuchi *et al.* 1992, 1993) has emerged as a technological leap forward based on the revolutionary method of PCR. The advantage of real-time PCR (RT-PCR), in comparison with conventional PCR technology, is that it allows better determination of the starting amount of DNA in the sample (before the amplification started) and "real-time" monitoring of PCR products as they are generated (Valasek and Repa 2005). Real-time PCR is also called "quantitative PCR" because of the ability to measure and quantify specific targeted nucleic acid sequences in the sample at the beginning of PCR, then monitor and record the amplification progress using fluorescent technology, and quantifying the amount of PCR product at the completion of the PCR (Valasek and Repa 2005). RT-PCR also helps in monitoring the specificity of produced PCR product, contamination and PCR inhibition (Willerslev *et al.* 2004). Since real-time PCR machines became commercially available in 1996, more and more researchers have started to use this advanced PCR technology (e.g. Yue and Orban (2001) on fresh and preserved fish scales; Zsikla *et al.* (2004) on formalin preserved clinical samples; Swango *et al.* (2006) for the assessment of DNA degradation in forensic samples; Gilbert *et al.* (2007b) for assessing the quality of formalin-arDNA, and so on). It provides much more information about samples and PCRs, and allows easier interpretation of confusing results than with the "traditional" PCR technique (Valasek and Repa 2005; Gilbert *et al.* 2007b).

## 1.3.3. DNA sequencing

The genetic material of each living organism possesses sequences of its nucleotide building blocks that are uniquely and specifically present only in its own species. These distinctive variations in a given genomic region make it possible to identify with precision what species (Floyd *et al.* 2002; Hajibabaei *et al.* 2006; Karaiskou *et al.* 2007) and often which group or particular member of that species it is (Gibson and Muse 2002). DNA sequencing is considered as the molecular technique with the highest level of resolution in comparison to other molecular techniques (Schlotterer 2004). It also provides results with easy cross-study comparisons and with already existing data repositories (e.g. database of the National Centre for Biotechnology Information; NCBI). This method was initially time-consuming

and expensive, but recent advances in sequencing technology permit automated, high-throughput sequencing and fast sequence analysis of many DNA fragments and for many individuals at once (Gibson and Muse 2002; Schlotterer 2004). Consequently, in the last few years, there have been more and more genome-scale sequencing projects that study the whole genome of organisms (Clark *et al.* 2004; Boore *et al.* 2005; Peck *et al.* 2005).

The amplified DNA can be sequenced either by direct sequencing of PCR products (preferably) or by sequencing of cloned PCR fragments (Brown 2001). Clones might not always faithfully represent the sequence of the original template DNA molecule because of errors occasionally introduced by *Taq* polymerase during the amplification (see section 1.3.2.1; p 18). If an erroneous PCR-DNA copy is cloned and sequenced, it will produce an incorrect sequence of the template because a clone is the exact multiplication of only one PCR-DNA copy which can carry a correct or incorrect sequence (Brown 2001). However, this characteristic has found its application for the verification of the authenticity of sequences extracted from samples with degraded and damaged DNA by cloning of PCR products and sequencing of multiple clones derived from more than one independent amplification (Hebsgaard *et al.* 2005; Wandeler *et al.* 2007).

Direct sequencing of a PCR product and the correct sequence of the template will usually not be affected by the errors introduced by *Taq* polymerase because of the errors that are, in the majority, distributed randomly (Brown 2001). For every molecule that has an error at a particular nucleotide position, there will be many other molecules with the correct sequence and, in this case, the error rate is usually insignificant. The exception could be only in cases if PCR amplification begins with only one, or very few DNA molecules, and if *Taq* polymerase introduced errors that occurred in the first cycle of the PCR (Hofreiter *et al.* 2001b). Although direct sequencing is faster, involves less labour and is more reliable, there are experiments where sequencing of cloned PCR products is unavoidable (for instance, sequencing of amplified RAPD fragments).

Finding species-specific DNA sequences for species identification ("signature sequences" and/or DNA "barcodes") is important for molecular taxonomy and in assessing biodiversity (Schander and Willassen 2005), but accumulation of DNA sequences from different regions of different species and comparison of sequences between species is important in establishing homology of DNA sequences and finding highly conserved genome regions and loci (Gibson and Muse 2002). The correct sequence and correct sequence alignment are of crucial importance to comparative molecular biology and genome science. Sequence

alignment is basically a procedure by which it is possible to recognise and describe a potential homology among nucleotide, or amino acid positions (Wheeler 1994; Gibson and Muse 2002). Software now exists automatically to align, edit and compare sequences (e.g. SEQUENCHER – Gene Codes Corporation; CLUSTAL - Higgins and Sharp (1988); MALIGN - Wheeler and Gladstein (1993); GeneJockey II - Taylor (1993); etc.), which is a great help in speed and accuracy of sequence analyses. Aligning and comparing sequences of investigated organisms, looking for sequences that are similar to those already identified in other genomes by using comparison algorithms such as BLAST (Basic Local Alignment Search Tool), and applying gene-finding software that recognises DNA features that are associated with genes, such as open reading frames (ORFs), as well as searching for homologous proteins in other species, are steps towards the development of genome science that studies the structure, content and evolution of genomes (Gibson and Muse 2002).

Assessing and retrieving sequence information from genetically uncharacterised species are of special importance for building up a "full picture" of a particular group of organisms and their relationships with other species and groups of organisms (Boyle *et al.* 2004; Schander and Willassen 2005; Wandeler *et al.* 2007). Single-nucleotide polymorphisms (SNPs), i.e. SNP variations, along with variation in repetitive sequences (such as microsatellites) are important in characterising DNA sequence diversity and for inferring relationship between individuals, populations, species, and in evolutionary studies of the history of a species (Schlotterer 2004; Gilbert *et al.* 2007c; Wandeler *et al.* 2007).

### 1.3.4. Molecular methods based on a difference in electrophoretic mobility

Many methods have been developed to assay variation in DNA sequences. The majority of them rely on differences in the DNA sequences being reflected in differences in the chemical properties of the molecule (Schlotterer 2004). Some of methods that detect these changes and detect sequence variation through a difference in electrophoretic mobility are (from: Schlotterer 2004): denaturing gradient gel electrophoresis (DGGE), temperature gradient gel electrophoresis, single-strand confirmation analysis (SSCA), heteroduplex analysis, single cell gel electrophoresis (SCGE). Mentioned methods might be potentially useful to adapt and apply in detecting DNA damage and changes (e.g. cross-linking, fragmentation) caused and induced by chemicals used for the preservation of archival specimens (see Carter (2003) for the use of DGGE and heteroduplex analysis for estimating DNA damage introduced by different preservation treatments of museum specimens).

### 1.3.5. Molecular markers

The existing marker systems and molecular methods can be categorised depending on what they can reveal and which part of the genome or kinds of loci are under investigation (Liu and Cordes 2004; Schlotterer 2004). Genetic markers can be distinguished as nuclear and organellar (mitochondrial and plastid), depending on the part of the genome under investigation. Single locus and multiple loci marker systems can be recognised on the basis of genome coverage, i.e. number of loci detected. Also, it is possible to divide molecular genetic marker systems on the basis of the requirement for prior DNA sequence information. Molecular markers could be derived from known regions of genome, or they could be anonymous from unknown parts of genome. It is possible to make different categorisations and classifications of molecular marker systems, but the basic principle of all molecular genetic methods is to employ inherited, discrete and stable markers for identifying genotypes that characterise individuals, populations or species (Schlotterer 2004).

#### *1.3.5.1. RAPD-PCR markers*

Randomly amplified polymorphic DNA markers consist of relatively short DNA fragments (about 200-3000 base pairs long), amplified via PCR by small arbitrary primers (usually 10 bases in length) (Williams *et al.* 1990). Because RAPD-PCR primers are not designed to amplify a specific target sequence, the amplified loci could be in any part of genome (Williams *et al.* 1990; Clark and Lanigan 1993). RAPDs are dominant markers, shown by the presence or absence of a band on a gel, enabling genotypic data to be obtained simultaneously at many loci (Hill and Weir 2004). RAPD loci carry the advantage that there is no need for prior nucleotide sequence data for the taxa under study. This makes RAPD-PCR methodology suitable for investigations of genetically uncharacterised species where specific sequence information is lacking.

The technical simplicity, cost-effectiveness, and speed in producing RAPD profiles are also advantages of the RAPD-PCR methodology. In many taxa RAPD-PCR can easily generate hundreds of independent markers (Yang *et al.* 2006). Many of the RAPD markers may exhibit substantial variation in evolutionary rates because different loci (genomic regions) may have different mutation and recombination rates through evolutionary history and events (Grosberg *et al.* 1996; Isabel *et al.* 1999). They are also useful for genome mapping

(Krutovskii *et al.* 1998) and detecting transposable elements (Dioh *et al.* 2000). Generated RAPDs give the opportunity of detecting population-, species-, genus- or family specific markers and can be applied for the rapid identification of organisms at different levels and to ascertain genetic diversity (Hadrys *et al.* 1992; Bielawski and Pumo 1997; Sommerfeldt and Bishop 1999; Bardakci 2001; Chiari and Sodre 2001; Pearson *et al.* 2002; Almeida *et al.* 2003; Birmeta *et al.* 2004; Wan and Fang 2003; DeLaat *et al.* 2005). Generally, RAPD markers provide high resolution of genotype distribution in natural populations (Jacobson *et al.* 1993; Baratti *et al.* 1999, 2003; Prioli *et al.* 2002).

RAPD markers can vary in several ways. The most common is the presence of a band in some individuals and its absence in others (Hill and Weir 2004). The usual cause of band absence is a mutation in one or both of the primer sites, resulting in the inability of the primer to hybridise (Hunt and Page 1992). Length variation can also occur if, in some individuals, a primer site is present between two original primer sites (Routman and Cheverud 1994). RAPD amplification products can be broadly classified into two groups: variable (polymorphic) or constant (non-polymorphic) (Hadrys *et al.* 1992).

Application of RAPD technology for obtaining taxon-specific information and developing taxon-specific PCR primers is demonstrated in many studies. For instance, Miyata *et al.* (1996) revealed and sequenced species-specific RAPD fragments for bacterial strains known to cause disease in salmonid fishes. Kambhapati *et al.* (1992) successfully applied RAPD-PCR technology for species identification and differentiation of conspecific populations of *Aedes* mosquito. Similarly, its use has been demonstrated in studies on *Paramecium* (Skotarczak *et al.* 2004), a crustacean – *Gammarus* species (Costa *et al.* 2004), tilapia fishes (Ahmed *et al.* 2004), and so on. In the Garcia *et al.* (1994) study, the RAPD-PCR methodology proved to be useful in isolating microsatellites, estimating genetic diversity within and between populations, identifying family-specific markers, and mapping loci responsible for economically important traits in penaeid shrimp.

It is possible to convert a RAPD multilocus genetic marker into a more informative single-locus marker system (Hadrys *et al.* 1992). Single-locus RAPD markers can be subcloned and then used as specific probes for conventional RFLP analysis (Hadrys *et al.* 1992; Fani *et al.* 1993), or for developing taxon-specific primers, STS (Sequence Tagged Site) and SCAR (Sequence Characterised Amplified Region) markers as sequence-specific markers (e.g. Hadrys *et al.* 1992; Paran and Michelmore 1993; Barret *et al.* 1998; Agusti *et al.* 2000; Scheef *et al.* 2003; Bailey *et al.* 2004; Das *et al.* 2005; Hughes *et al.* 2006). These

approaches can circumvent the major limitation in the application of RAPD methodology and is possible to produce more specific and reproducible results, i.e. results that are less sensitive to changes in PCR reaction conditions (Scheef *et al.* 2003; Yue *et al.* 2003). Converted, they can be used as a multiplex PCR assay for fast detection and discrimination of species (Fernandez *et al.* 2003).

#### 1.3.5.1.1. Disadvantages and problems with the RAPD-PCR methodology

There are many advantages of RAPD-PCR methodology, but there are also some serious limitations and problems related to its application. First, the inconsistency and non-reproducibility of RAPD-PCR profiles (RAPD-PCR fragments) put the reliability of this methodology in question (Skroch and Nienhuis 1995; Jones *et al.* 1997; Rajput *et al.* 2006). Some other problems related to the practical use of RAPD (and MAAP techniques in general) are:

(a) Amplification fragments of apparently the same molecular weight, generated with single arbitrary primers, can result from the targeting of different genetic loci (van de Zande and Bijlsma 1995),

(b) Different amplification products can result from the targeting of the same genetic locus. In other words, parts of the same sequences could be represented in more than one fragment in RAPD patterns (Smith and Williams 1994).

These complicate interpretation and cross-gel comparisons. In order to reduce these problems, Southern blotting, or cloning and sequencing of corresponding RAPD fragments, should be applied (Hadrys *et al.* 1992; Clark and Lanigan 1993; Rabouam *et al.* 1999). Sometimes RFLP analysis of RAPD-PCR fragments could be a fast way to investigate homology of comigrating RAPD fragments (Rieseberg 1996), but this application has limitations. It might be successful for large fragments if sites for applied restriction enzymes are present, but RFLP is not suitable for RAPD-PCR fragments smaller in size. The reasons are: (1) small fragments are less likely to carry restriction sites, and (2) the gel resolution becomes critical for the correct classification of very small fragments.

Another disadvantage of the RAPD technique is the production of multi-banded profiles with bands that encompass a variety of intensities (ranging from strong to faint). This makes repeatability and comparison of RAPD-PCR experiments difficult (Foissner *et al.* 2001; Skotarczak *et al.* 2004). There are many factors that can affect reproducibility of RAPD-

PCR fragments and the intensity of bands, such as: quality and quantity of a PCR-DNA template, commercial brand of thermostable DNA polymerase, type of PCR thermal cycler and PCR conditions applied, and so on (see Schierwater and Ender 1993; Jones *et al.* 1997). Hallden *et al.* (1996) are of the opinion that competition in the amplification of RAPD fragments (competitive priming) is a major source of genotyping errors in RAPD analysis.

A serious disadvantage of RAPDs is the possibility of generating artifactual RAPD-PCR fragments (Ellsworth *et al.* 1993; Rabouam *et al.* 1999). Rabouam *et al.* (1999) suggest that the sources of RAPD artifacts might be: (1) the presence of commensal organisms (such as viruses, bacteria, mycoplasma or protozoa), i.e. contaminating DNA, and (2) fragment rearrangements during PCR amplification (Fig. 1).

Rabouam *et al.* (1999) suggested that there are probably three types of interactions responsible for the synthesis of artifactual RAPD fragments (Fig. 1):

(1) intrastrand interactions,

(2) interstrand interactions, and

(3) nested primer annealing.

*Intrastrand interactions* are due to short inverted repeats within the RAPD fragments leading to the formation of dyadic intrastrand interactions. *Interstrand interactions* may involve the annealing of similar termini between RAPD fragments. They may also involve the annealing of short repeated and similar sequences between RAPD fragments. The last type is the *nested primer site binding* that generates shorter and larger fragments.

Rabouam *et al.* (1999) investigated the formation and the frequency of the artifactual RAPD bands (fragments) by applying Southern blotting for detecting erroneous, i.e. genuine RAPD-PCR fragments (bands). They reported 50-75% RAPD-PCR fragments as being artifactual. In their opinion, it was, therefore, impossible to estimate the degree of DNA polymorphism by RAPD analysis. Rabouam *et al.* (1999) are of the opinion that the number of these artifactual RAPD fragments probably differs greatly between genomic DNA sources and depends on the RAPD primers applied; that is, that their study and findings on *Calonectris diomedea* (Cory`s shearwater, Aves) and *Haemonchus contortus* (Nematode) may not, therefore, establish general rules for RAPD studies for all taxa under investigation.

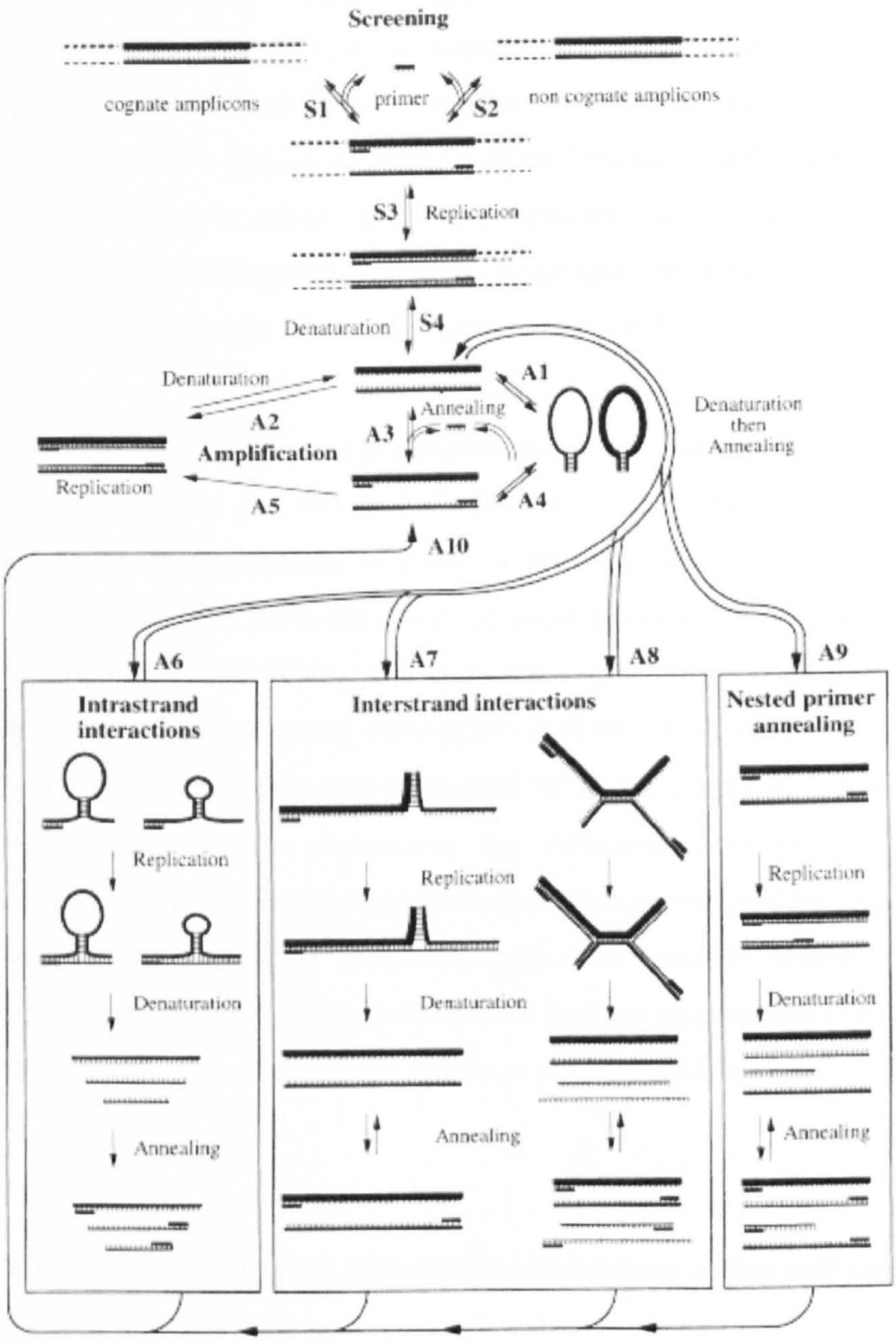



**Fig. 1** – Rearrangements of fragments in the RAPD-PCR products as a possible source for production of artifactual RAPD fragments (from: Rabouam *et al.* (1999) – modified model of Caetano-Anolles *et al.* (1992))

Other serious problems are related to the gel reading and scoring of RAPD fragments (bands), their statistical analyses and correct interpretation of results (Hadrys *et al.* 1992; Garcia *et al.* 1994; Hadrys and Siva-Jothy 1994; Isabel *et al.* 1999; Nybom 2004). Many researchers count only strong bands that are reproducible in most RAPD-PCR amplifications in independent runs (Cook *et al.* 1998; Gouin *et al.* 2001; Collins *et al.* 2003; Grazziotin and Echeverrigaray 2005; Guerra-Garcia *et al.* 2006), but this criterion is not always applicable and/or adequate (see Jones *et al.* 1997).

Issues discussed above put RAPD methodology in the centre of a long-lasting debate about its reliability and the use of RAPDs as informative molecular markers. Some researchers claim high reproducibility and reliability of results (Penner *et al.* 1993; Gomes *et al.* 1998; Lougheed *et al.* 2000; Klinbunga *et al.* 2001; Segarra-Moragues and Catalan 2003; Guerra-Garcia *et al.* 2006), whereas others doubt the reproducibility of RAPD-PCR amplified fragments and patterns across experiments in the same and different laboratories (e.g. Paran and Michelmore 1993; Ellsworth *et al.* 1993; Jones *et al.* 1997).

There are interesting observations and a comprehensive statement produced by Schlotterer and Pemberton (1994) about the use of PCR-RAPD methodology: "Anyone with a PCR machine is able to use this technique and after a few attempts, will be able to produce a polymorphic pattern from the genomic DNA of most organisms. More experienced PCR users doubt the simplicity of RAPDs, as it is well-known that every PCR amplification is extremely dependent on the purity and concentration of the DNA as well as on the reaction conditions. Even if the researcher is able to control the above parameters, another drawback of RAPDs will remain: Due to dominance the difference between homozygotes and heteroyzgotes is not visible. Discrimination between homozygotes and heteroyzgotes cannot be neglected for population analysis. Especially critical for paternity testing with RAPDs, it is the occurrence of non-paternal bands that can lead to false conclusions (Scott *et al.* 1992). Last, but not least, there are no data on the mutation rate for RAPDs."

#### 1.3.5.1.2. <u>The usefulness of the RAPD-PCR methodology – why still use it?</u>

Despite the many disadvantages of RAPD-PCR methodology and numerous reservations regarding the reproducibility, homology, and statistical analyses of data (see for a review Grosberg *et al.* 1996), several highly attractive features (such as: a sensitivity, applicability in cases when sequence data or any other molecular data are absent for the taxon under investigation, theoretically limitless number of genetic markers, tool for developing other classes of molecular markers, etc.) make RAPD methodology still highly desirable for molecular investigations (Hadrys and Siva-Jothy 1994; Beebee *et al.* 1999; Collins *et al.* 2003; Ali *et al.* 2004; Yang *et al.* 2006).

In experiments with some organisms and biological materials, the mentioned problems are not so expressed and RAPD methodology is still being successfully used for traditional

measurement of polymorphism and differences in genetic structure of populations (Lougheed *et al.* 2000; Theodorakis *et al.* 2006), subspecies (Echeverrigaray *et al.* 2001) and species (Isoda *et al.* 2000; Collins *et al.* 2003; Grazziotin and Echeverrigaray 2005), or establishing kinship (Cao and Card 1997; Edwards *et al.* 2002; Shikano and Taniguchi 2003; Wallace 2006). Some researchers are of the opinion that if it is used carefully, its results are reliable (Collins *et al.* 2003; Pearson *et al.* 2002; Vilatersana *et al.* 2005). The reliability of RAPD-PCR technology, including a technique of scoring bands and an application of statistical analysis for the measurement of polymorphism was tried to be improved by some researchers (Hansen *et al.* 1998; Collins *et al.* 2003; Ritland 2005), but more researchers are now inclined to explore new applications of this methodology (for instance, Abe *et al.* (1998a,b, 2000, 2005) in detecting retrotransposable elements; Cook *et al.* (1998) in detecting polyploidy; Adams *et al.* (1999) and Tayutivutikul *et al.* (2003) in testing the quality of DNA extracts and PCR usability for difficult biological samples; Guha and Kashyap (2006) in identification of unknown samples in legal medicine using so-called exclusion approach) and to develop other classes of molecular markers by using RAPDs as an initial approach (see Behura *et al.* 1999; de Kloet 2001; Scheef *et al.* 2003; Yang *et al.* 2006).

RAPDs might be useful in conservation and environmental research through screening archival collections for revealing the genetic structure of populations that lived in a particular space and time (Koh *et al.* 1997; De Wolf *et al.* 2004). RAPD fingerprint analyses based on band numbers (gain/loss of bands) and RAPD band intensity might yield diagnostic genetic information (De Wolf *et al.* 2004; Enan 2006). There are reports that RAPD methodology can facilitate these applications on fresh/frozen and short period ethanol-preserved material (De Wolf *et al.* 2004; Wu *et al.* 2005; Enan 2006; Theodorakis *et al.* 2006), but only further research on differently preserved specimens can demonstrate if the same approach (with adequate modifications) is also applicable to archival collections with specimens stored for longer periods and subjected to different fixation/preservation treatments.

Similarly, RAPDs have also found application in biomedical science in the evaluation of general rearrangements of the genome associated with tumorigenesis, i.e. in detecting DNA alterations shown by "band gain" or "band loss", decreased/increased intensity of RAPD bands, or banding shifts (Ong *et al.* 1998; Maeda *et al.* 1999; Luceri *et al.* 2000; Luo *et al.* 2003).

Evidently, the potential of the RAPD method is not only in its application for polymorphism analysis, but also in applying a RAPD-based strategy as an initial approach for developing other classes of molecular markers and/or as a diagnostic approach for detecting the changes in genetic structure.

#### 1.3.5.1.3. RAPD application to preserved specimens

It is expected that the use of degraded/damaged DNA for RAPD polymorphism profiling multiplies the mentioned problems and is much more complicated than its application on "good" DNA (see Eckerman and Walsh 1997; Carvalho and Vieira 2000; Oliveira *et al.* 2002; Siwoski *et al.* 2002; Boyle *et al.* 2004). However, damage to the DNA may not interfere with analysis, raising the possibility of using this methodology on specimens that have not been sampled or preserved under conditions ideal for DNA integrity (see A`Hara *et al.* 1998; Anami *et al.* 2000; Wang *et al.* 2005). In experiments where other molecular methodologies and molecular markers are not applicable for different reasons, RAPD-PCR methodology might be the only molecular approach available for initial molecular investigations (see the results of this study).

### *1.3.5.2. DNA microsatellite markers*

Microsatellite markers have found application in many molecular investigations and proved to be useful for addressing different biological questions (Schlotterer 2000; Liu *et al.* 2006; Selkoe and Toonen 2006). It is possible that primers designed for one species might be applicable to another, usually closely related species (see Hogan and May 2002; Pandey *et al.* 2004), but this depends on the stability of the repeat loci over evolutionary time and the rates of evolution of the sequences flanking the repeats (Ashley and Dow 1994; Brohede and Ellegren 1999). Generally, new microsatellite primers need to be developed for almost every species which may be problematic for non-model organisms and taxa with unstudied genomes (Schlotterer 2004), especially if available only/mostly as samples with degraded/damaged DNA. This is because of a requirement for substantial sequence information on the investigated taxon – for screening genomic libraries, or to have a large amount of a high-molecular-weight DNA for probes that can be used for developing informative microsatellite markers. But, there are alternative ways of isolating

microsatellites by analysing sequences of anonymous nDNA regions (for example, using RAPD-PCR fragments which are known to be enriched in microsatellites; Ender *et al.* 1996).

#### 1.3.5.2.1. Isolation of microsatellites and its possible application on samples with degraded DNA

The possibility of detecting microsatellites by sequencing RAPD-PCR fragments and then screening them for tandem repeats (Garcia *et al.* 1994; Adcock *et al.* 2000; Liu *et al.* 2006) and/or applying microsatellite probes on RAPD-PCR fragments (Ender *et al.* 1996; Lunt *et al.* 1999; Iyengar *et al.* 2000; Seyoum *et al.* 2007) give an opportunity of isolating microsatellites and applying this marker on genetically uncharacterised species (lacking sequence information), including samples with degraded/damaged DNA (Jakobsdottir *et al.* 2006). This approach is usually more efficient for a large RAPD-PCR fragment, because there is a higher chance that larger DNA fragments contain microsatellite regions (with its flanking sequences) than smaller RAPD-DNA fragments (Ender *et al.* 1996), but smaller fragments can also reveal microsatellites (Jakobsdottir *et al.* 2006). The possibility of finding microsatellites and other repetitive elements in RAPDs (DNA RAPD-PCR fragments) might lie in the fact that RAPDs most likely reflect the repetitive sequences of large genomic blocks; that is, consist of repetitive sequences (Ghany and Zaki 2003; Ribeiro *et al.* 2004; Yang *et al.* 2006).

Application of microsatellites to museum specimens and other samples with degraded/damaged DNA is possible, although degraded DNA might cause errors in PCRs and allelic dropout, that can lead to erroneous genotyping (see Taberlet *et al.* 1999; Gill 2001; Lucchini *et al.* 2002; Jovanovic *et al.* (2003); Piggott and Taylor 2003; Sefc *et al.* 2003; Rohland *et al.* 2004; Wisely *et al.* 2004; Tie *et al.* 2005; Watts *et al.* 2007).

## *1.4. The importance of museum collections in biological studies*

Natural history museum collections have been assembled and maintained over the course of more than 200 years (Lane 1996). Historically, all these collections were established in order to preserve morphological features of organisms and to be used for morphological analysis. The traditional view of a museum as just being a specimen repository changed a long time ago. Collections have been actively used in different aspects of research (Huber 1998), but for the purpose of systematic research in particular (Rainbow 2001). Information from natural history collections about the diversity, taxonomy, systematics, historical distributions of species worldwide and their phylogenetic relationship are becoming increasingly important as a heritage which needs to be maintained, understood and protected, especially in a period with evident decline and extinction of present species (Thomas 1994; Wirgin *et al.* 1997; Dayton 2003; Goldstein and DeSalle 2003; Austin and Melville 2006; Fong *et al.* 2007), as well as in an era of finding and identifying many new species (Hebert *et al.* 2004). This means that museum collections begin to have a new dimension where the genetic material of different organisms might be stored and be used for different purposes (see Baker 1994; Pettitt 1994; Winker 2004; Austin and Melville 2006). Natural history collections are an irreplaceable database of information and a historical record of changes in populations and biodiversity during the past century or more (Thomas *et al.* 1990; Wayne and Jenks 1991; Krajewski *et al.* 1992; Roy *et al.* 1994; Thomas 1994; Wirgin *et al.* 1997; Bouzat *et al.* 1998; Shaffer *et al.* 1998; Doukakis *et al.* 2000; Pergams *et al.* 2003; Graham *et al.* 2004; Parham *et al.* 2004; Harper *et al.* 2006). For some groups of organisms and parts of the world, museum collections might soon represent the only record of biodiversity.

Another fact is that collecting samples for a particular project can be expensive and difficult to achieve - for instance, in the case of deep-sea studies, or Antarctic research. The use of already collected and stored samples makes such research projects more feasible to conduct. For example, Rex *et al.* (2005b) and Zardus *et al.* (2006) point out the importance of the existence of faunal deep-sea archival collections (below 200 m) for studying spatial and temporal aspects of population differentiation and deep-sea biodiversity. There are many reasons to use museum specimens in a study rather than going out and collecting anew. These mostly fall under the headings of temporal, spatial, political and financial reasons, and combinations of these (Thomas 1994; Schander and Halanych 2003; Suarez and Tsutsui 2004). The use of archival world collections prevents unnecessary destruction of many

individuals of different species used for study by various research teams, although extensive destruction of museum specimens (for molecular and other "destructive methodologies") should be also prevented (Wisley *et al.* 2004; Gilbert 2007d). This is possible to achieve by placing all data on voucher specimens and a list of related publications on the web page of the relevant museum and/or into dedicated database repositories (Graham *et al.* 2004).

Thomas (1994) made suggestions about using museum collections for molecular studies and sharing obtained information on samples from collections of natural history museums:

- to make available material from museum collections for scientific investigations, but research laboratories should make written proposals and a brief outline of the study plan before sampling is permitted;
- when sampling, the smallest possible amount of material should be removed, and this should be taken, where possible, from tissues that contain few characters of morphological interest;
- aliquots of DNA extracts and PCR products remaining from a study of museum materials to be returned to the museum for permanent storage, but these are not available to others until the researcher has published on them, or indicated that publication will not be forthcoming; sequence data or database accession numbers should be provided – the intention being to minimise the need for repeated sampling, except for the purpose of result re-examination and re-validation;
- museum and curators must be properly acknowledged in all publications derived from museum specimens, so that it will be possible to demonstrate the productivity of museums and scientific institutions.

Many of these recommendations have already been adopted by the majority of museums. Integration of data on museum specimens around the world and their easier access is currently in progress (Graham *et al.* 2004). Molecular data from museum specimens aid in understanding of past and present, and should help in predicting the future of many processes related to flora and fauna across the globe (Wandeler *et al.* 2007).

Until recently, morphological analysis was the only approach that scientists applied in their research on museum specimens, but now, molecular biology is seen as further enhancing the value of collections and offers the potential for exploring a huge amount of genetic information stored in museum collections (Rivers and Ardren 1998; Rogers 2001; Bhadury *et al.* 2006a). There are a few possible limitations to conducting research projects solely on museum specimens such as: the number of individuals (especially of big animals) is not

always sufficient for particular investigations (e.g. population studies), difficulties in obtaining museums` permission for tissue sampling of museum specimens, difficulties in extracting sufficient DNA in quantity and quality required for molecular investigations (Schander and Halanych 2003; Austin and Melville 2006; Tang 2006).

### 1.4.1. Preserved specimens in molecular investigations

Technical advances in molecular biology are increasing the range of biological material amenable to molecular analysis, but this field still remains difficult and challenging (Hajibabaei *et al.* 2005; Wandeler *et al.* 2007). It is necessary to develop and standardise efficient techniques for DNA retrieval from preserved specimens, especially from fluid-preserved collections (Austin and Melville 2006; Tang 2006; Gilbert *et al.* 2007b,c; Skage and Schander 2007). There is a need to establish some guidelines, similar to what has been done in the field of ancient DNA and bioarchaeology (Yang *et al.* 1997a,b; Cooper and Poinar 2000; Machugh *et al.* 2000; O`Rourke *et al.* 2000; Gilbert 2003a; Paabo *et al.* 2004; Cipollaro *et al.* 2005; Gilbert *et al.* 2005a). Many biological, physical and chemical factors that affect ancient DNA quality are also relevant for specimens from natural history collections, such as endogenous nuclease activity and hydrolytic damage (Wandeler *et al.* 2007). Also, many recommendations and approaches adopted in bioarchaeology are also applicable in molecular work with preserved specimens, but some specifics are related exclusively to preserved archival collections (see Austin and Melville 2006; Tang 2006; Gilbert 2007b; Karaiskou *et al.* 2007; Skage and Schander 2007; Wandeler *et al.* 2007). We have a general understanding about the (possible) damage of DNA extracted from preserved specimens, but we still lack the specific knowledge related to differently preserved specimens of the same and different collection age - the extent of the DNA damage induced by different fixatives and preservation solutions/chemicals, types of DNA damage and their distribution in DNA molecules/sequences, the possibilities of reversing and repairing some types of damage (Tang 2006; Skage and Schander 2007; Wandeler *et al.* 2007).

The type of fixation and the conditions of preservation as extrinsic factors, as well as the type of tissue, species and size of an organism as intrinsic factors have an effect on the survival and extractability of DNA in preserved specimens (Cooper 1994; Thomas 1994; Tayutivutikul *et al.* 2003; Boore *et al.* 2005). The aim of fixation is to inhibit decay and to produce covalently bonded cross-linked macromolecules that have the effect of holding tissue together and retaining a resemblance of the organism's appearance in life (Criscuolo

1994; Thomas 1994; Martin 2004). Preservation has the function of keeping tissues and organisms in good morphological shape over long periods of time without degradation (Carter 2003; Martin 2004). Good morphological preservation of museum specimens does not mean that quantities of intact DNA are present for molecular studies (Thomas 1994; Carter 2003). While protein-protein, DNA-protein and DNA-interstrand cross-links are desirable for morphology, they are undesirable for molecular investigations because cross-links inhibit the polymerases used to replicate DNA in PCR, thereby preventing amplification (O`Leary *et al.* 1994; Thomas 1994; Dorris 1999; Karaiskou *et al.* 2007; Skage and Schander 2007).

There are also many other factors (such as: specimen collection and handling of the samples, kinds and purity of chemicals used in preservation, storage temperature and maintaining collections) that are important and associated with the success of DNA investigations on preserved specimens (Prendini *et al.* 2002; Boyle *et al.* 2004; Tang 2006; Gilbert 2007d).

#### *1.4.1.1. The importance of sample collecting and fixation/preservation treatments*

Both collection of the organisms in the field and fixation/preservation treatments are important for later success in molecular investigations (Prendini *et al.* 2002).

##### 1.4.1.1.1. Sample collection

The method of collecting specimens in the field is important for the quality and quantity of DNA in general, but it is significant for archival specimens in particular because chemical preservation and a prolonged storage periods cause further DNA degradation (Schander and Halanych 2003; Miething *et al.* 2006; Ferrer *et al.* 2007). As is well known, the processes of degradation and fragmentation of DNA start relatively soon after an organism dies because of DNase activity and autolytic decay (Lindahl 1993a,b; Brown 1999; Briggs 1999; Vass 2001; Yoder *et al.* 2006). Some researchers suggest that the most DNA damage, DNA fragmentation and degradation occur in the first couple of hours after an organism dies (Perry *et al.* 1988; Paabo 1993).

The environmental conditions (temperature, humidity, chemical surroundings, etc.) under which a dead organism was kept before fixation (freezing or drying out) are also important

for DNA "survival" and its usability for molecular investigations (Schander and Halanych 2003). Degradation and fragmentation of DNA might be accelerated if optimal conditions for nuclease activity and decay are present (Prendini *et al.* 2002). However, the collection of animals from particular habitats might be more difficult to carry out in a desirable way than others because of problems in controlling time between death of the organism and placing it into a fixative solution (or at subzero temperature). A good example of this is collecting deep-sea fish specimens. A length of time between when a fish dies and when dead fish are brought to the surface might vary very much - it could be from half an hour to a couple of hours. This timing cannot be controlled by the collector because it depends on the depth from which fish are caught, fishing equipment and technique applied, and on the sensitivity of a particular species to temperature changes and changes in hydrostatic pressure (Bailey *et al.* 2002). However, time when the dead animals are on the surface until they are placed into fixative can be controlled by a collector and it is important to do it as soon as possible. All together a record should be made of "total time" of death for deep-sea organisms (post-mortem delay in animal preservation). Another aspect, which might also need to be considered in collecting marine organisms in general, is whether the presence of sodium chloride (NaCl) and other compounds in the sea and in the marine organisms has any effect in accelerating the activity of enzymes and other natural processes in a dead animal that later might affect the quantity and quality of the DNA.

All this information, together with information on applied fixation/preservation treatments, is important for the successful use of museum and other archival collections as valuable sources of DNA. For example, some collectors and curators applied fixative by boiling formalin in order to achieve better penetration of formalin into the organism (Waren 1983), a process which might have devastating consequences for DNA in the preserved organism. Often all this information is not available (especially not for earlier collections), or recorded data are inaccurate (see Austin *et al.* 2003; Stuart *et al.* 2006). The accurate recording of all these data (including any chemicals used as narcotics and for relaxing the organisms, and/or chemical treatments for retaining colour and unchanging dimensions of specimen) is important (see, for example, Dean and Ballard (2001) on the effects of insect killing methods on mtDNA yield and PCR success). This missing precise information is an additional problem that complicates attempts for DNA protocol standardisations (Tang 2006). A molecular biologist often has no choice in selecting specimens that might be collected and preserved in a way that are more suitable for molecular work (i.e. with supposedly better DNA preservation). Bucklin and Allen (2004) suggest that checking the pH of storage/preservation solution might be helpful in this regard. In other words, low pH

(< 4) is likely to be an indicator of extensive DNA degradation and fragmentation, i.e. of specimens of a low suitability for molecular work.

#### 1.4.1.1.2. Fixation and preservation of archival specimens

Fluid-preserved and dried samples are the most common types of preserved specimens in museum collections. Formalin fixation (usually 4%-10% of formalin) and then ethanol preservation (usually in 70%-80% of ethanol, i.e. IMS) are the norm in museum fluid collections, but other chemicals, such as methanol, isopropanol, acetone, chloroform, acetic acid, mercuric chloride, various arsenic- and lead-based compounds have been used for some organisms (Carter 2003). It seems that dried specimens (treated with chemicals or not) are easier to apply in molecular investigations than fluid-preserved ones (see Rohland *et al.* 2004; Irestedt *et al.* 2006). Formalin is considered as the most damaging fixation/preservation procedure for molecular work (Silberman and Walsh 1992; Schander and Halanych 2003; Tang 2006).

FORMALIN – Formaldehyde

**Formaldehyde** (HCHO) is a gas, the aldehyde produced by partial oxidation of methanol (methyl alcohol) – Dioni (2002). As a gas it is used in many chemical industrial processes, but in biological applications it is used by dissolving it in water. Formaldehyde reacts with water to form an equilibrium mixture of water, formaldehyde, and formaldehyde hydrate (Dioni 2002). **Formalin** is an aqueous solution of formaldehyde (Kiernan 2000; Dioni 2002). Commercial formalin contains 37-40% of formaldehyde and 60-63% of water (by weight) (Kiernan 2000; Dioni 2002). Formaldehyde is self-reactive, continuing to oxidize in aqueous solution producing formic acid, and in older solutions may even form a precipitate of paraformaldehyde (solid polymerized formaldehyde) (Kiernan 2000; Dioni 2002). It is common to add to formalin 10% or so of methanol as a stabilizer. Formalin solution thus really contains formaldehyde, paraformaldehyde, formic acid and methanol (Dioni 2002).

So formalin has a target composition but only a detailed analysis can determine what the true concentration of any component is in a given batch of a given age. Another problem related to the use of formalin for preserving biological samples is the fact that formalin is almost certainly acidic and this could impair its performance as a fixative or preservative (pH is usually 2.8-4; Schander and Halanych 2003). Acidity and hydrolytic attacks on DNA

are well known as extremely damaging, causing strands breakage, nucleotide modifications and lack of bases at various sites in the DNA sequence (Koshiba *et al.* 1993; Lindahl 1993b; Douglas and Rogers 1998; Hoss *et al.* 1996; Poinar 2002; Bonin *et al.* 2003; Paabo *et al.* 2004). The rates of depurination and strand cleavage are increased at acidic pH and at elevated temperatures (Poinar 2002). It is the general belief that by using neutral buffered formalin (made by adding some buffers to the 10% formalin solution – usually sodium phosphate for a pH 7.0-7.2, or sodium borate (borax) for a pH about 7.7), it is possible to improve the performance of formalin and balance the changes that could be occurring, i.e. to stabilise the behaviour of the fixative (see Carter 1997).

There is also a lot of confusion in the use of the terms "formalin" and "formaldehyde" and their real meaning in published literature (Dioni 2002). Authors tend to use these terms interchangeably, as well as using interchangeably the terms formalin and formaldehyde solutions of concentrations 4% and 10%. Sometimes, it is very difficult exactly to understand which concentration, in real terms, they are referring to. Both words have their own and very different meanings of great importance for morphological fixation and preservation (Dioni 2002; Gershwin 2002), and for molecular work in particular.

## STEEDMAN'S SOLUTION

Steedman's solution as a preservation/storage solution is mostly used for herpetological and ichthyological museum collections (Carter 2003; Martin 2004), but also for preservation of plankton (Steedman 1976; Bottger-Schnack *et al.* 2001), crustaceans and molluscs (Carter 2003). This preservative solution, besides formalin, also contains other chemicals and additives, such as a propylene glycol and propylene phenoxetol, and in order to improve its preservation, anti-bacterial, fungicidal and anti-oxidant chemicals (Steedman 1976). There are no published molecular data on specimens preserved in Steedman's solution, but it is expected to show poor DNA preservation and potentially degradive effects on DNA because of the presence of unbuffered formalin and propylene phenoxetol ($C_6H_{12}O_2$) (Carter 2003). The negative effect of propylene phenoxetol on morphology of long-term preserved fish specimens was discussed and reported by Crimmen (1989). Nakanishi *et al.* (1969) also reported that proteins are more sensitive to denaturation in the presence of phenoxetol than without it at room temperature. The effect of phenoxetol on DNA is completely unknown, but it is assumed that it could only have negative effects on the integrity of DNA and PCR usability (Carter 2003).

This study is one of the first to explore the possibility of molecular work on museum specimens preserved/stored in Steedman's solution. Carter (2003) attempted to extract DNA from Steedman's preserved museum specimens, but applied a DNA extraction protocol (CTAB [hexadecyltrimethylammonium bromide] method) that did not yield PCR amplifiable DNA for the tested molecular markers (mitochondrial 16S and ribosomal 18S), and the DNA was also unsuitable for restriction endonuclease screening.

The success in obtaining molecular data from fluid-preserved specimens varies from collection to collection (Wandeler *et al.* 2007). For example, France and Kocher (1996) recorded different amplifying success with 10% buffered formalin-fixed, ethanol-preserved (FFEP) deep-sea crustaceans that came from collections which claimed to be fixed and preserved in the same way as successful ones. Some of these collections were donated, some were collected by other researchers, and some specimens were collected from other localities than successful ones. France and Kocher (1996) also found problems in a yield of PCR products if DNA was extracted from specimens that were extensively manipulated following transfer to ethanol for sorting samples by sex under a microscope. They also were not able to produce PCR amplifiable DNA if it had been extracted from formalin-fixed, formalin-preserved (FFFP) samples stored for several months, or if specimens were preserved in glutaraldehyde. They successfully extracted DNA only if specimens were fixed in 10% buffered formalin (with 90% sea water) for at least 24 hours, transferred into 80% ethanol for at least 24 hours, and finally stored in a fresh solution of 80% ethanol. The DNA extracted from these specimens was successfully used for PCR amplifications of mitochondrial 16S and CO I genes. This is a well-described example indicating a collection effect (including possible differences between localities and the sample handling during collection), fixation of samples and post-fixation manipulation on the DNA extractability and PCR usability.

Zardus *et al.* (2006) reported a similar situation about usability (i.e. non-usability) of formalin-fixed specimens of a deep-sea bivalve that came from different collections, i.e. different stations and localities. Organisms that lived in a particular area might absorb specific chemicals during their lifetime (see Forlin *et al.* (1996) and Sole *et al.* (2001) for fish) that might have different effects on DNA extractability and PCR usability. Researchers reported inconsistency in producing PCR amplifiable DNA not only from different collections and/or localities, but also between specimens from the same collection (e.g. Quattro *et al.* 2001). The differences between specimens of one particular collection could happen if individuals were not handled in the same way during collection, fixation and/or

preservation (Prendini *et al.* 2002). Small discrepancies in fixation and preservation process are not so important for morphological investigations, but in extracting PCR-amplifiable DNA, it seems that the smallest variation in collecting and preserving specimens might cause dramatic differences in the ability to use preserved specimens for DNA studies (Giannella *et al.* 1997; Schander and Halanych 2003; Miething *et al.* 2006). The importance of each step in the fixation procedure for DNA studies is clearly presented in the paper of Moore *et al.* (2002).

Until recently, the only focus of museum curators was to find the best chemical treatments (often mixing different chemicals) to maintain good gross morphological preservation of specimens and to retain morphological features that are important as taxonomic characters for species identification (Treasurer 1992; Shields and Carlson 1996; Thibault-Botha and Bowen 2004). Even maintaining collections to preserve morphological features for long periods is not an easy task (see Tucker and Chester 1984; Stephenson and Riley 1995; Black and Dodson 2003; Haussermann 2004; Gagliano *et al.* 2006). Adding different chemicals in order to retain authentic pigmentation and coloration, and/or to maintain other morphological and histological features of the preserved organism might significantly change the pH and other characteristics of the fixative/preservative solutions (Taylor 1981; Deveney and Whittington 2001), thereby affecting the possibility of DNA recovery from these archival collections (Prendini *et al.* 2002; Schander and Halanych 2003).

Understanding the importance of maintaining gross morphology with recognisable taxonomic characters of specimens for morphological analyses (Paradis *et al.* 2007), but also the difficulties (and, in particular cases, inability) in using chemically preserved museum specimens in molecular investigations, some researchers have suggested that two forms of sample storage should be applied for new acquisitions (Dessauer and Hafner 1984; Criscuolo 1994; Thomas 1994; Prendini *et al.* 2002). One is a traditional, chemical preservation, and another is cryopreservation and/or pure ethanol preservation of tissue/specimens for maintaining DNA in good condition (see Corthals and DeSalle 2005). Cryopreservation is "a golden rule" for keeping DNA in good condition for a long time (Dessauer *et al.* 1996; Ferrer *et al.* 2007), but ethanol can also be used for a short-term storage (probably not longer than five years for the majority of organisms; Hajibabaei *et al.* 2005). The potential problem with ethanol preservation is that ethanol might be unsuitable for DNA recovery from some species (see Reiss *et al.* (1995) for beetles; Dorris (1999) for nematodes), and another reason is that ethanol might oxidise and/or become acidic unless it is regularly replaced, diluted with water and/or buffered (Criscuolo 1994; Hajibabaei *et al.*

2005). Currently, effort is also directed towards finding a fixation/preservation solution that will be suitable for both, good morphological and good DNA preservation for longer periods (van Dam 2003; Rey *et al.* 2004; Delfour *et al.* 2006; Mandrioli *et al.* 2006; Molin and van Dam 2006; Yoder *et al.* 2006).

#### 1.4.1.1.2.1. *DNA damage induced by formalin*

The difficulties in using formalin-fixed material for molecular investigations were reported by many researchers in biological, forensic and medical science (e.g. Romero *et al.* 1997; Herniou *et al.* 1998; Whittier *et al.* 1999; Cawkwell and Quirke 2000; Legrand *et al.* 2002; Liborio *et al.* 2005; Chakraborty *et al.* 2006; Tang 2006; Skage and Schander 2007). Although there is still no agreed uniform view on the kinds and degree of DNA damage caused by formalin, most researchers are of the opinion that formalin degrades DNA, causing considerable DNA damage and also causes DNA inaccessibility from preserved specimens (Brown 1999; Legrand *et al.* 2002; Srinivasan *et al.* 2002; Bonin *et al.* 2003; Martin 2004; Melton *et al.* 2006; Zardus *et al.* 2006). Broadly speaking, there are four possible types of DNA damage: denaturation, cross-linking, strand breakage and chemical modification (Brown 1999). There are suggestions that deleterious effects on DNA are promoted by the first step in formalin-fixed tissue/specimen processing (e.g. Shiozawa *et al.* 1992; France and Kocher 1996; Rait *et al.* 2006) and that prolonged storage only continues to cause further damage to DNA (Pavelic *et al.* 1996). Many investigators are of the opinion that preservation/storage in ethanol, after formalin fixation (24-48 hours), is much less damaging for DNA than preservation/storage in formalin (Diaz-Viloria *et al.* 2005). Further investigations and confirmation of the exact facts about formalin-induced DNA damages is important because most museums (and other collections) contain preserved specimens usually stored in alcohol, but originally fixed in formalin.

Thus far, there have been suggestions that:

- Formaldehyde causes rapid denaturation of DNA followed by strand breakage and also could cause chemical modifications of nucleotides (Vachot and Monnerot 1996).

- Others suggest that the main interaction of DNA and formalin is the formation of cross-links between protein and DNA (Koshiba *et al.* 1993), which causes little if any substantial damage to the DNA polymers (Haselkorn and Doty 1961). Karlsen *et al.* (1994) suggests that the DNA itself remains non-fragmented, but possibly altered in structure.

- Furthermore, some researchers are of the opinion that the main problem is not the action of formaldehyde but the possibility that formaldehyde oxidizes to formic acid causing chemical modification to DNA (Taylor 1977). Koshiba *et al.* (1993) is of the opinion that if unbuffered formalin is kept at room temperature for long periods, the acidity of formalin will increase considerably, causing hydrolytic attacks on DNA and extensive DNA degradation.

These are very different perceptions about possible DNA damage caused by formalin. Some of these views may be out of date, but are presented as historically evolving perceptions related to this issue. Contemporary knowledge has no doubt that formaldehyde produces inter- and intramolecular cross-links in proteins, DNA-interstrand cross-links and cross-links between proteins and nucleic acids (Chaw *et al.* 1980; Solomon and Varshavsky 1985; Chang and Loew 1994; Orlando *et al.* 1997; Douglas and Rogers 1998; Kiernan 2000; Quievryn and Zhitkovich 2000; Nitta *et al.* 2002; Liu *et al.* 2006; Rait *et al.* 2004, 2006; Barker *et al.* 2005a,b; Yamashita and Okada 2005; Tang 2006; Yamashita 2007). Toth and Biggin (2000) even reported the specificity of protein-DNA formaldehyde crosslinking, i.e. that protein and DNA crosslink to discrete regions of the genes/DNA sequences under the influence of the formaldehyde, as well as that formaldehyde may not crosslink some proteins to DNA. This might be of significance in studying particular genes/genome regions of formalin-fixed specimens from archival collections.

Beside the specificity of proteins that crosslink to DNA, Barker *et al.* (2005b) also reported about the differences in mechanisms of forming DNA-protein crosslinks (DPCs) and their different chemical stability (Fig. 2). According to Barker *et al.* (2005b), proteins can become crosslinked to DNA:

- directly through oxidative free radical mechanism (mainly irreversible DPCs; these DPCs are generated within a short distances of each other in the DNA and are usually in association with "locally multiply damaged sites" and "clustered lesions" that involve DNA single strand breaks, double strand breaks, base damage or base loss, DNA-DNA crosslinks), or
- they can be crosslinked indirectly through a chemical linker (the majority of aldehyde-induced DPCs are unstable and can be reversed), or
- through coordination with a metal atom (can induce different chemical types of crosslinks – stable and unstable, covalent and non-covalent bounding of proteins to DNA, what depend on kind of metal ions involved in crosslinking).

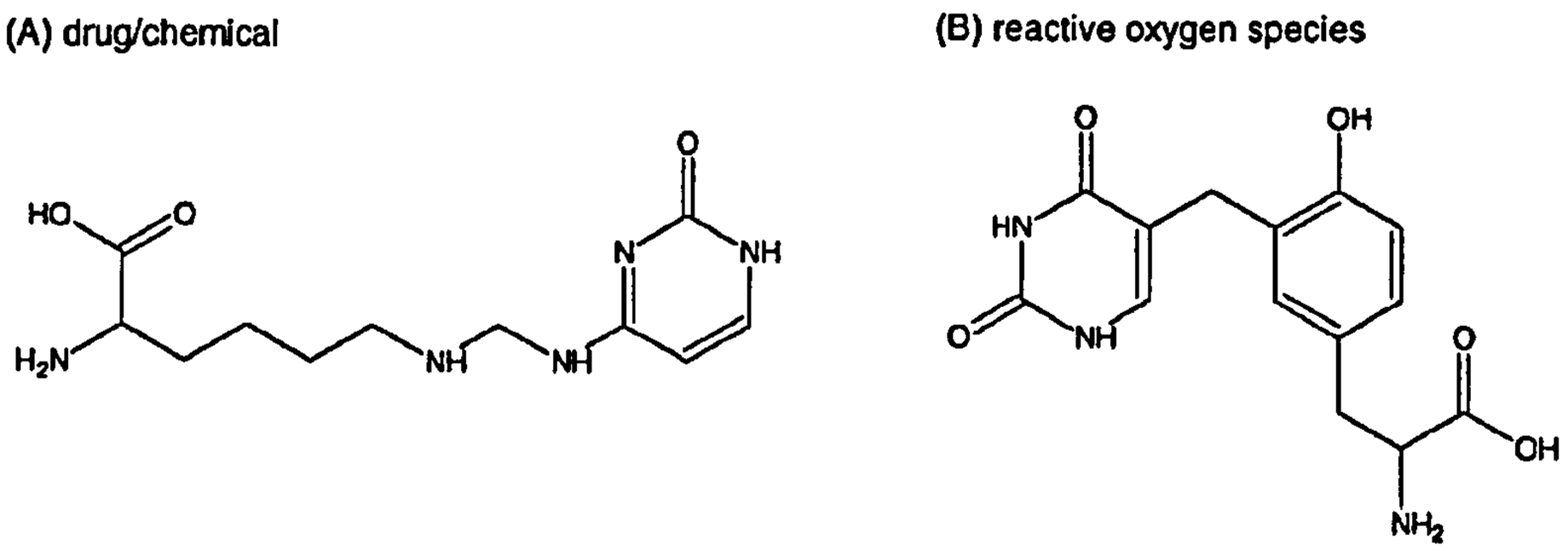



**Fig. 2 - Crosslink structures.** A schematic representation of two of the chemistries by which proteins may become crosslinked to DNA. **(A)** A formaldehyde induced crosslink between cytosine and lysine **(B)** A radiation-induced crosslink between thymine and tyrosine. (From: Barker *et al.* 2005b).

Most probably, DNA extracted from formalin-fixed tissues/specimens contain a combination of different types of damage considering the possible effects of the death of organism (i.e. delay in tissue/organism fixation), effect of the formaldehyde (including the effects of the potential acidity of formalin solution used for fixation/preservation), the presence of metal ions and other chemicals in a body of a particular organism and type of tissue (as natural constituents of body/tissue and from surroundings in which the organism lived, including environmental pollutants at particular localities).

Dorris (1999) observed and discussed reversing the effects of formalin fixation and ethanol preservation in nematodes from aspects of the suitability of extracted DNA for molecular investigations. Dorris (1999) suggests that formaldehyde causes some structural changes to DNA chains by denaturation and progressively destroying DNA by hydromethylation, or irreversibly cross-linking formaldehyde with proteins via thiols, phenolic groups and terminal amino groups. However, in Dorris's (1999) opinion, for material recently fixed in formalin, the process of fixation appears to be reversible at least for PCR products up to 1000 bp. The inability to amplify greater sized products suggests that formaldehyde is affecting some mechanism of irreversible fixation. Dorris (1999) stated that his procedure and results provide strong evidence that formaldehyde fixation is reversible to some extent and does not reduce the fidelity of the DNA sequence to the degree previously assumed.

Dubeau *et al.* (1986) (on pathology samples), France and Kocher (1996) (on formalin-fixed deep-sea crustaceans), Shedlock *et al.* (1997) (on fish specimens), Chase *et al.* (1998a), Quattro *et al.* (2001), Boyle *et al.* (2004) and Zardus *et al.* (2006) (on deep-sea molluscs) also did not find any, or no significant, alterations of PCR-DNA sequenced fragments derived from formalin-fixed archival samples (in comparison with fresh/frozen and/or

ethanol-preserved ones). But, Williams *et al.* (1999) and Quach *et al.* (2004) reported a high frequency of sequence alterations due to formalin fixation of archival specimens. De Giorgi *et al.* (1994) also found infidelity of PCR-amplified DNA fragments induced by formalin. Estimates so far suggest that more than 20% of sequences derived from formalin-preserved specimens show one or more errors (Wandeler *et al.* 2007).

Many researchers are of the opinion that types of damage and the extent of damage to the DNA, as well as the degree of cross-linking and its reversibility, are related to the length of tissue exposure to the formalin and to the appropriateness of the formalin buffering (Dubeau *et al.* 1986; Crisan and Mattson 1993; Dorris 1999; Bucklin and Allen 2004). However, Douglas and Rogers (1998) reported that DNA of plant material treated with 10% formalin – acidic or neutral (pH 3.0, or pH 7.0) was so badly damaged in both cases that DNA could not be amplified. Silberman and Walsh (1992) suggested that the use of formalin as a preservative for plankton collections should be avoided, as this degrades DNA to the extent that it can no longer be PCR-amplified. In their studies, they did not have particular problems with extraction of DNA and PCR amplification from specimens of the same preservation age that were ethanol preserved.

The majority of researchers consider ethanol to be a good medium for maintaining DNA in good condition (Simmons 1991; Post *et al.* 1993; O`Leary *et al.* 1994; Whittier *et al.* 1999; King and Porter 2004; Linville *et al.* 2004; Chakraborty *et al.* 2006). There are also suggestions that DMSO-NaCl (Dawson *et al.* 1998; Kilpatrick 2002), and 40% isopropyl alcohol (Shiozawa *et al.* 1992) are good preservatives for specimens intended to be used for molecular studies, but the best results are obtained from untreated dry specimens (Rohland *et al.* 2004) and frozen tissue (see the review of Prendini *et al.* (2002) about the most appropriate fixative/preservative solutions for molecular work).

#### *1.4.1.2. DNA extractions, improvements of protocols, and problems*

The general belief is that museum specimens contain a very small amount of amplifiable endogenous DNA and precautions must be taken at all stages during extraction and amplification processes to minimise opportunities for contamination with extraneous DNA (Thomas 1994; Bucklin and Allen 2004; Bhadury *et al.* 2006a; Wandeler *et al.* 2007). Museum specimens have been handled by humans and are often processed or stored in groups where individuals are in contact with one another, or even fixation/preservation fluid

has been exchanged between different containers (Wandeler *et al.* 2007). This might cause cross-contamination. For example, Chase *et al.* (1998a) and Schander and Halanych (2003) found human DNA contamination in some mollusc DNA extracts isolated from formalin-fixed, ethanol-stored museum specimens.

Another problem in work with preserved specimens is that impurities in DNA extracts can inhibit or reduce the efficiency of PCR amplifications (An and Fleming 1991; Reiss and Rutz 1999; also see the section 1.3.2.1.1; p 22). The selection of an optimal DNA extraction method and purification steps are crucial if analysis of DNA extracted from archival specimens is to be successful (Herniou *et al.* 1998; Kullmann *et al.* 1998; Poljak *et al.* 2000; Yue and Orban 2001; Legrand *et al.* 2002).

Several classes of DNA extraction protocols and purification procedures are in common use:

1) Digestion with proteinase-K followed by phenol-chloroform extraction (e.g. Shiozawa *et al.* 1992; Shedlock *et al.* 1997; Mandrioli *et al.* 2006);
2) Use of a chelating resin such as Chelex 100 (BioRad), which binds metal ions, combined with proteinase-K digestion (e.g., Walsh *et al.* 1991; Eckerman and Walsh 1997; Coombs *et al.* 1999; Kirby and Reid 2001; Perez *et al.* 2005) and silica (Yue and Orban 2001);
3) Guanidinium thiocyanate extraction (Hammond *et al.* 1996; Whittier *et al.* 1999);
4) Selective binding of DNA to a matrix, such as glass, silica, or diatomaceous earth (e.g., Boom *et al.* 1990; Cano and Poinar 1993; Evison *et al.* 1997; Bertozzini *et al.* 2005; Ivanova *et al.* 2006).
5) A CTAB (hexadecyltrimethylammonium bromide) extraction method (Yang *et al.* 1997a; Doukakis *et al.* 2000; Carter 2003; Fuchs *et al.* 2005).
6) Commercially available DNA extraction kits such as: "Qiagen QIAamp DNA minikit" (Coombs *et al.* 1999; Bernstein *et al.* 2002; Karaiskou *et al.* 2007), "Qiagen DNase tissue extraction kit" (Beebee *et al.* 1998; Austin and Melville 2006; Irestedt *et al.* 2006), "Promega – The Wizard Genomic DNA Purification Kit" (DeLaat *et al.* 2005);
7) Combination of particular DNA extraction steps from developed protocols and commercial kits (e.g. Chase *et al.* 1998a; Coura *et al.* 2005; Bhadury *et al.* 2006a).

For fluid-preserved specimens, it is necessary to apply appropriate pre-extraction treatments prior to DNA extraction in order to remove residual preservation solution from the tissues.

Some of commonly used treatments (mainly for formalin-fixed tissues) are listed below:

- Tissue washing with TE buffer (Shiozawa *et al.* 1992; Wirgin *et al.* 1997; Li *et al.* 2000; Klanten *et al.* 2003).
- Tissue washing with glycine-based buffer (1xGTE washing buffer; Shedlock *et al.* 1997), or a combination of 1xGTE buffer, ethanol and water (Hasbun *et al.* 2005).
- Wu *et al.* (2002) applied a 15 minute pre-incubation of the formalin-fixed tissue at 98°C in Tris-based buffer.
- Shi *et al.* (2002, 2004) applied tissue heating at an alkaline pH: 100°C - 120°C at pH 9-12, demonstrating a greater yield of DNA and reduced level of PCR inhibitors in DNA extracts.
- Critical point drying using a series of 30%-100% ethanol washes for a gradual dehydration and removal formaldehyde from specimens (Fang *et al.* 2002).
- Tissue washing in phosphate-buffered saline (PBS) buffer – recommended as a part of protocol for DNA extraction from preserved tissue by the "Qiagen" and "Promega" for the use with their commercially available kits.

There are some studies that have investigated relationships between DNA extractability/DNA extraction protocols/PCR usability in conjunction with a particular preservative treatment (Dawson *et al.* 1998; Boyle *et al.* 2004; Chakraborty *et al.* 2006; Mandrioli *et al.* 2006; Gilbert *et al.* 2007b) and with the exposure lengths of a tissue/organism to a particular fixative/preservation treatment (Rey *et al.* 2004; Bhadury *et al.* 2005, 2006a,b; Ferrer *et al.* 2007; Skage and Schander 2007). Literature available on the subject brought to light some discrepancies. For example, Shedlock *et al.* (1997) described successful amplifications of fish mtDNA (16S for 570 bp and cytochrome b for 470 bp) from formalin-fixed, ethanol-stored specimens of different storage periods (including one specimen preserved for 85 years), whereas Chakraborty *et al.* (2006) reported an inability to extract PCR amplifiable mtDNA (16S and 12S for fragments of about 600 bp and 400 bp) using the same DNA extraction protocol (and five other) for formalin-fixed fish specimens stored for 3-4 years in 70% ethanol and/or formalin. These suggest the need for much more investigations in order to better understand preserved specimens from the bimolecular aspect and the possibility of their applications to molecular-genetic studies.

Recently, some more comprehensive results were published in this field (e.g. Vachot and Monnerot 1996; Brown 1999; Prendini *et al.* 2002; Carter 2003; Schander and Halanych 2003; King and Porter 2004; Hajibabaei *et al.* 2005; Tang 2006; Mtambo *et al.* 2006; Gilbert *et al.* 2007b,c,d; Skage and Schander 2007; Wandeler *et al.* 2007), but this kind of

information is still well behind guidelines and overviews presented in forensic science (e.g. Haglund and Sorg 1997, 2002; Hoff-Olsen *et al.* 1999; Rudin and Inman 2002; Edson *et al.* 2004; Buckleton *et al.* 2005) and in research with ancient DNA (e.g. Herrmann and Hummel 1994; Wayne *et al.* 1999; Cooper and Poinar 2000; Brown 2001; Hummel 2003; Gilbert 2003a; Poinar 2003; Paabo *et al.* 2004; Cipollaro *et al.* 2005; Gilbert *et al.* 2005a; Willerslev and Cooper 2005; Binladen *et al.* 2006; Hansen *et al.* 2006; Poinar *et al.* 2006). Also, much more testing and recommendations are provided by researchers in biomedical science (Koshiba *et al.* 1993; Giannella *et al.* 1997; Coombs *et al.* 1999; Cawkwell and Quirke 2000; Sato *et al.* 2001; James *et al.* 2002; Shi *et al.* 2002, 2004; Srinivasan *et al.* 2002; Inadome and Noguchi 2003; DeVries *et al.* 2005; Miething *et al.* 2006; Ferrer *et al.* 2007) than by researchers in fundamental biological science.

Molecular investigations on preserved specimens require a multidisciplinary approach; it is time-consuming and requires laboratory-intensive work, with the outcome that much less genetic information is obtainable (at present) than with the same amount of time and laboratory work with DNA from good sources (fresh/frozen and short-term ethanol preserved material) (Wandeler *et al.* 2007). In addition, validation of results and analyses of data obtained from preserved specimens (and other specimens with degraded DNA) usually require a specific approach, which, again, is not fully developed and standardised (see Gilbert *et al.* 2007b; Wandeler *et al.* 2007). All these make molecular work with preserved specimens very difficult and unpredictable in comparison with "good" DNA. These are the probable reasons that molecular studies on preserved specimens were neglected for a long time and that researchers in biological science have tended to use good sources of DNA for their research.

However, it seems that global barcoding project with its intention to include specimens from natural history collections activated and accelerated the development of this field of fundamental biological science. Very recently, more comprehensive reviews and research on preserved specimens with useful observations, recommendations and guide-points for the further research have been published (Tang 2006; Gilbert *et al.* 2007b,c,d; Skage and Schander 2007; Wandeler *et al.* 2007). Hopefully, the realisation of the possibility of extracting useful and reliable molecular data from archival specimens, and the scientific significance of these data (DNA sequences in particular) as valuable records from the past, will accelerate further progress in the field.

### *1.4.1.3. Some examples of molecular studies on preserved specimens*

Most studies have used high copy number sequences, such as nuclear ribosomal genes and mitochondrial genes that improve the chances of finding amplifiable DNA (e.g. Dawson *et al.* 1998; Doukakis *et al.* 2000; Bhadury *et al.* 2005, 2006a; Zardus *et al.* 2006). The possibility of finding single copy nuclear genes that are amplifiable in a given volume of material is extremely difficult and unlikely with old and preserved specimens (Thomas 1994; Greenwood 2001; Rogers 2001; Stuart *et al.* 2006).

One example of studies that involved museum specimens is the study of Doukakis *et al.* (2000) who tried to re-examine the species identification of three museum sturgeon specimens (two alcohol preserved and one stuffed specimen) and to clarify the biogeography of two sturgeon species (*Acipenser sturio* and *A. naccarii*). Two specimens were originally described as *Acipenser sturio* L. on the basis of morphometric analyses, but later the same two specimens were identified as *A. naccarii*. Garrido-Ramos *et al.* (1997) suggest that the Adriatic sturgeon, *A. naccarii*, also existed historically in Spain, but only the proper identification of these sturgeon museum specimens would provide the necessary evidence, because *A. naccarii* is no longer found in Spain. Doukakis *et al.* (2000) applied molecular investigations on the mitochondrial genome: portions of the cytochrome b, NADH5, and 12S genes, and portions of the non-coding central region of mtDNA. Three separate laboratories investigated the same preserved fish specimens, but they were not able to identify species with absolute confidence because of the inability to obtain consistent results. This example indicates that preservation and long storage of museum specimens make reliable DNA extraction and PCR amplification difficult; that is, it caused difficulties to apply effectively molecular investigations in order to obtain unambiguous results. However, this example emphasises:

(1) the importance of the existence of *Acipenser* museum specimens (with fishes of populations/species that are may be no longer present at the original habitat) available for retrospective studies from the collection,

(2) the importance of further developments in molecular biology - development of DNA extraction protocols, molecular techniques, approaches and markers suitable for application to museum specimens and other difficult samples with degraded DNA.

Another example is that of confusing results related to differences in PCR usability (for 18S ribosomal gene) of different formalin-fixed collections of marine nematodes (collections aged from two weeks to up to 20 years) in association with different experimental procedures (Bhadury *et al.* 2005, 2006a). The use of unbuffered formalin gave very poor amplification results for ~400-bp size product after only 9 days of the specimens' storage - PCR yield became low and inconsistent (Bhadury *et al.* 2005). They were unable to generate amplification product after 11-15 days. The period of 11 or 15 days depended on the kind of thermostable DNA polymerase that was used in PCR experiments. *Pfx* DNA polymerase (Invitrogen) generated PCR products for specimens stored in formalin for up to 15 days, whereas *Taq* DNA polymerase (Promega) generated PCR products only for specimens stored in formalin for up to 11 days. However, by using different collections, different extraction protocols (hot lysis protocol) and by designing an internal primer for amplifying the product size of ~345 bp from the same region of the 18S gene, they successfully amplified targeted DNA of all tested nematodes (25 specimens) that were fixed and stored in formalin for up to 20 years (Bhadury *et al.* 2006a), even using Promega *Taq* DNA polymerase which did not exhibit such good PCR effectiveness in the previous experiments on DNA of nematodes that were preserved in formalin for less then two weeks (Bhadury *et al.* 2005). This is an example of very different results being obtained after the application of different extraction protocols and experimental strategy, and by using different formalin-fixed collections of different storage times. It will be useful to investigate if the successful laboratory procedure from the experiments in Bhadury *et al.* (2006a) will be also successful for the nematode collections investigated in Bhadury *et al.* (2005) in order to establish if the main reason for differences in the obtained results was different preservation effect (the use of different formalin collections associated with the different effects of collecting and preserving of organisms), or if it was DNA extraction protocol/experimental strategy that was the main reason for differences in results in these two experiments.

Molecular data, especially sequence information, derived from museum specimens and other archival collections are very valuable and can be potentially applicable in different biological fields for many different studies (Thomas *et al.* 1997; Schander and Willassen 2005; Wandeler *et al.* 2007). However, the understanding of preserved specimens from the aspect of their applicability in DNA studies is a field that needs much more clarification and information than is available at the present.

## 1.5. ***Problems with species with uncharacterised sequences***

The biological investigations that involve species with uncharacterised DNA sequences and genetically uncharacterised groups of organisms are exposed to the limitations of the use of available molecular tools and techniques, that is, it might be time-consuming until appropriate markers are found (Schander and Willassen 2005). The application of universal PCR specific primers (mitochondrial and nuclear), or developed specific primers for a particular group(s) of organisms (Kocher *et al.* 1989; Folmer *et al.* 1994), might be limited in studying new animal systems (species with unstudied genomes) because the flanking sequences and priming sites may differ so much in comparison with successfully tested groups of organisms as to cause inability of primers annealing (failure of PCR amplifications).

If good sources of DNA are available, testing the applicability of already existing PCR primers and/or isolation of molecular markers from unstudied genomes should not be so problematic (Liu *et al.* 2006). But, if genetically uncharacterised species are studied only as samples with degraded DNA (archival preserved specimens, bioarchaeological material, non-invasive samples, etc.), finding appropriate markers for their study might be extremely difficult (Wandeler *et al.* 2007). Inability to apply universal and other specific PCR primers might not be only related to inability of primer annealing due to inadequate priming sites (differences in base composition of flanking sequences), but also due to DNA damage of relevant DNA regions (missing and/or chemically altered parts of the targeted sequences, cross-linking) and/or the presence of PCR inhibitors in DNA extracts, producing false negative and/or confusing PCR results (Tang 2006; Gilbert *et al.* 2007d). Some researchers have reported successful development of taxon-specific primers from preserved organisms with unstudied genomes by using DNA region-specific universal primers as an initial approach (Boyle *et al.* 2004), indicating that this strategy might be applicable if universal primers generate PCR products of sufficient yield and quality for sequencing. However, if universal and other specific PCR do not produce successful PCR amplifications, employing random priming methodologies is probably the best approach in conducting the research on such species. This approach proved to be useful in many studies (see the section 1.3.5.1; p 28), including this study on *Nezumia*.

## *1.6. Problems with populations and species from remote areas*

Investigations that consider remote populations/species have, often, a very difficult assignment. Species from remote populations are usually genetically uncharacterised, that is, without available information on their DNA sequences (Koenemann *et al.* 2006). Even basic information about distribution and population structuring, speciation and evolution are often insufficient - if they exist at all (Gage 2004; Madurell *et al.* 2004; Brandt 2005). If species/populations from remote localities are of a low density (as usually they are), if they have low abundance and/or are endemic for a particular area(s), a study often has to be performed on a small number of individuals because of difficulties in collecting specimens in the field (Kearney and Stuart 2004).

One solution is the use of museum and other archival collections as part of a particular study in order to increase the number of individuals from a particular location (see Collins *et al.* (2002) in studying swamp eels; Kearney and Stuart (2004) in studying worm lizards). In some studies, archival collections and preserved specimens are the only source of samples for molecular and other investigations for organisms that come from remote areas (e.g., Boyle *et al.* (2004) and Zardus *et al.* (2006) in studying population structuring of a deep-sea bivalve). The deep sea, the poles, caves and deserts are typical examples of remote and difficult-to-sample localities where individuals of many species are difficult to collect fresh in sufficient number. The existence of archival collections is important and can significantly improve knowledge about organisms that live in such areas (Schander and Willassen 2005; Skage and Schander 2007).

## *1.7. Introduction to the marine environment including the deep sea*

Two-thirds of the earth's surface is covered by water, with the majority being ocean of more than 500 m depth (Creasey and Rogers 1999; Angel 2003). The deep sea is, therefore, important and the world's largest ecosystem with a mean depth of ~3800 m (Tyler 2003). The remoteness of the deep sea and the difficulties in its exploration has resulted in it being one of the least understood environments on Earth and with shortage of information on all kinds of organisms (see: Marshall 1979; Merrett and Haedrich 1997; Chase *et al.* 1998b; Creasey and Rogers 1999; Tyler 2003; Gage 2004; Etter *et al.* 2005; Rex *et al.* 2005b).

### 1.7.1. The marine environment

The sea is not a large "pond" which has a flat bottom and the same surroundings. The world ocean and the deep-sea floor have their own configuration and topography. The general pattern is one in which the rather uniformly shallow continental shelf (depth 0-200 m; proportion of total world ocean area: 7.5 %) gives way to a sea floor which inclines more steeply at the shelf break and down the continental slope (depth 200-2000 m; proportion of total world ocean area: 8.8 %) to a gently sloping area known as the continental rise with depth 1,500-5,000 m (Merrett and Haedrich 1997). At the foot of the continental rise lie the large areas of almost flat bottom, known as the abyssal plain (depth >4,000 m; proportion of total world ocean area: ~54 %, which covers most of the deep sea-floor; Merrett and Haedrich 1997). However, these generalised topographic regions have specific submarine landscapes that are variable and can be compared to terrestrial landscapes. Thus submarine canyons, hills and seamounts may punctuate the topography of the slope, rise or abyssal plain (in: Merrett and Haedrich 1997; Etter *et al.* 2005). Some of these physical oceanic "constructions" may go from very deep water up to the surface and range in width from a few hundred metres to several kilometres, or even several thousands of kilometres if mid-ocean ridges are included (from: Merrett and Haedrich 1997; Creasey and Rogers 1999).

Within the sea there are ecological islands of geographically restricted and physically distinct habitats isolated from each other by varying distances on scales of metres to thousands of kilometres. Such habitats include hydrothermal vents, hydrocarbon seeps, seamounts and cold-water coral reefs with specific flora and fauna compared to their surroundings (Probert 1999; Rogers 2000; Diekmann *et al.* 2006). Some of them are endemic species related to the particular region of the world ocean (Creasey and Rogers 1999; de Forges *et al.* 2000). All these factors, as potential barriers for some marine organisms, play an important role in influencing gene flow and population divergence (Creasey and Rogers 1999).

Studies of a number of different marine species have indicated that population genetic structuring can be high, despite apparently high dispersal capabilities and population mixing (e.g., Palumbi (1996b) on urchin; Miya and Nishida (1996) and Koslow *et al.* (1997) on fish; Benzie (1999) on starfish; Reusch *et al.* (2000) on marine plant). Knowlton *et al.* (1993) suspected that very rapid speciation may be occurring in widespread marine species, resulting in sibling taxa. The presence and number of marine cryptic species (deep sea in

particular) are issues that have been addressed (Quattro *et al.* 2001; Angel 2003; Etter *et al.* 1999, 2005; Moura *et al.* 2008). This can be reliably investigated primarily by the use of molecular methodologies, i.e. by combining molecular and morphological data, because morphological characters alone usually have insufficient resolving power for solving this problem in full (Colborn *et al.* 2001; Ward *et al.* 2005; Raupach and Wagele 2006). The introduction of molecular studies as a tool for population, systematic and taxonomic research has already provided better insight into biology, biography, ecology, interrelationships, gene flow and spatial boundaries of marine organisms (e.g. Creasey *et al.* 1997; Creasey and Rogers 1999; Morita 1999; Quattro *et al.* 2001; Miya *et al.* 2003; Wilson and Attia 2003; Williams and Reid 2004; Aboim *et al.* 2005; Hofmann *et al.* 2005; Peck *et al.* 2005; Rogers *et al.* 2006; Roques *et al.* 2006).

Distribution of larvae, juveniles and adults, and their migration patterns, could have an important impact in generating distinct subpopulations and discrete populations (Creasey and Rogers 1999; Cowen *et al.* 2000; Shaw *et al* 2004; Levin 2006) which subsequently may lead to the process of speciation. The relative importance to gene flow of passive egg or larval drift versus adult migration varies from species to species (Ward and Grewe 1995; Creasey and Rogers 1999; Cowen *et al.* 2000; Merrett 2003; Young 2003; Levin 2006). Species assemblages of larvae in particular regions of the sea could help in understanding this issue, but larval taxonomy is very difficult because of very few distinguishable morphological characters related to a particular stage of development, including juvenile stages (Merrett 2003; Sassa *et al.* 2004; Busby 2005). By developing species-specific molecular markers, the identification of eggs, larvae and juveniles should be much easier (Aranishi 2006; Karaiskou *et al.* 2007) and help in resolving issues about larval dispersal, distribution patterns and population structuring. Screening available archival collections of larvae/juveniles caught from different parts of the world can considerably facilitate in this regard (Karaiskou *et al.* 2007).

Generally, in the marine environment, it is possible to distinguish two patterns of diversity (Zardus *et al.* 2006): (1) geographical, that is, isolation by geographical distance; and
(2) bathymetric, depth-related divergence.

Zardus *et al.* (2006) suggest that genetic divergence in the deep-sea (for transect below 200 m and up to 3000 m in depth) is much greater among populations at different depths within the same basin than among those at similar depths but separated by long geographical distances (thousands of kilometres).

### *1.7.1.1. Deep-sea environments*

The division between shallow and deep water is defined as depth of 200 m because this transition represents a topographical and biogeographical transition that separates the two habitats (Gage and Tyler 1991; McClain *et al.* 2006). Exact definitions and boundaries of the bathyal and abyssal faunal provinces have, over the years, been subject to a number of approaches and interpretations (Menzies *et al.* 1973; Gage and Tyler 1991; Vinogradova 1997; Zezina 1997; Rex *et al.* 2005a). According to the bathymetric definition given by Gage and Tyler (1991), the bathyal zone refers to depths of 200-2000 m, whereas the abyssal zone is at depths of more than 2000 m below sea level. Zezina (1997) proposed that the bathyal faunal province ranged from 200 to 3000 m, whilst the abyssal environment comprised depths of 3000-6000 m, with the ultra-abyssal region occurring at depths of 6000 m or greater. Rex *et al.* (2005a) refers to ~ 4000 m in depth and 300 km from land as being the gradual transition between "bathyal" and "abyssal" habitats (source and sink habitats).

Bathyal and abyssal environments are characterised by relatively stable physical parameters: salinity, temperature, hydrostatic pressure and oxygen concentration (Menzies 1965; Tyler 1995; Creasey and Rogers 1999; Angel 2003; Etter *et al.* 2005). Of these four variables, temperature and salinity are relatively constant at any given location in the deep sea; at depths of 2000 m or greater, temperature varies from -1 to +4°C (on average being ~2°C; Thistle 2003) and salinity from 34.6% to 34.8% (Menzies 1965; Tyler 1995).

The special adaptations of species that inhabit different marine and deep-sea habitats are due to physical parameters and biotic conditions present in this kind of environment, which are likely to influence the behaviour and genetics of species that inhabit the particular niche (Palumbi 1994; Creasey and Rogers 1999; Peek *et al.* 2000; Thistle 2003; Etter *et al.* 2005). Within the deep-sea environment, vertical distributions of organisms may be limited in part by enzyme stability and biochemical adaptation to temperature and high pressure (Torres *et al.* 1979; Siebenaller and Somero 1989; Treberg *et al.* 2003). High pressure can result in modification of enzyme structure and changes in the volume of the protein (Somero 1982; Pennec *et al.* 1988; Cossins and Macdonald 1989; Siebenaller and Somero 1989; Morita 2003; Siebenaller 2003). This often includes molecular-level adaptations, such as modifications to the amino-acid sequence of an enzyme (Thistle 2003). Therefore, lack of appropriate adaptation of some species to specific conditions in deep-sea water could be the limiting factors in their bathymetric distribution.

We still know little about the number of fish species (Mason 2003) and other organisms that live in particular areas of the sea, especially in the deep sea (Koslow *et al.* 1997). For example, in cruises that took place in the New Caledonian region (Southwest Pacific Ocean) in the second half of 1990s, Iwamoto and Merrett found 63 species of grenadier fishes of which 19 were new species (Iwamoto and Merrett 1997; Merrett and Iwamoto 2000). This is a high proportion of new species found in a little-sampled area. The grenadier families (Macrouridae and Bathygadidae) are important constituents of the deep-sea demersal ichthyofauna (particularly on the mid and lower slope) and this represents around 15% of the species known to occur in the world ocean (Merrett and Haedrich 1997; Chiou *et al.* 2004a). New species of macrourid fishes (grenadiers) continue to be described (Iwamoto *et al.* 1999; McMillan 1999; Wilson 2001; Chiou *et al.* 2004a,b), as well as new species of other groups of fishes (e.g. Biscoito *et al.* 2001; Merrett and Moore 2005) and other organisms (e.g. Brenke *et al.* (2005) – new deep-sea isopod species).

It is clear that the deep sea contains many more species than previously thought, but our knowledge about the diversity and species richness in the deep sea, i.e. historical events and processes of speciation, is still very limited (Etter and Mullineaux 2001; Levin *et al.* 2001; Snelgrove and Smith 2002; Rex *et al.* 2005a,b; Kendall and Haedrich 2006). Gage (2004) and McClain *et al.* (2006) are of the opinion that much of the deep sea fauna died out in the mid-Cenozoic Era (due to significant changes in the temperature of the sea and a disoxyia/anoxia event) and were then replaced by shallow-water and coastal immigrants.

Recent evidence suggests that the most heterogeneous community of the deep sea is at bathyal depths between 200 to 4000 m and then gradually declines from 4000 to 6000 m (Etter *et al.* 2005; Olabarria 2005; Rex *et al.* 2005a). This might be related to oxygen-minimum zones (Rogers 2000; Angel 2003; Helly and Levin 2004), availability of food, dispersal capability of adults and larvae, selective gradients, various barriers to gene flow and historical events that have affected population differentiations and speciation (Etter *et al.* 2005). Rex *et al.* (2005a) proposed a source-sink hypothesis suggesting that many abyssal populations are maintained by immigration from the bathyal zone, and because of that, the abyssal plain is considered to have little potential for evolutionary divergence (see also Holt *et al.* 2003; Etter *et al.* 2005; Rex *et al.* 2005b). A low density of many abyssal populations suggests that they could not be reproductively viable and are vulnerable to local extinction (Etter *et al.* 2005; Rex *et al.* 2005a,b). McClain *et al.* (2006), who proposed body-size evolution in the deep sea, are of the opinion that abyssal organisms represent a unique

case in which source-sink dynamics lead to sink populations dominated by smaller individuals (see also Rex *et al.* 2005a).

Currently, small-scale regional and local studies of biodiversity, biogeography and population structuring of different species provide understanding of biodiversity, speciation and history of the deep sea as a whole, but limitation in sampling methodology, deficiencies in taxonomic knowledge and uneven geographic coverage are still serious obstacles in achieving comprehensive, global-scale knowledge about the deep sea and ocean (Gage 2004). A better use of museum and other archival collections provides the opportunity of expanding research from local to regional and global-scale investigations, including temporal aspects. There are few studies which included archival specimens for molecular investigation of genetic differentiation and population structuring of deep-sea organisms (France and Kocher 1996; Chase *et al.* 1998b; Quattro *et al.* 2001; Boyle *et al.* 2004; Bhadury *et al.* 2005, 2006a,b; Perez *et al.* 2005; Zardus *et al.* 2006), but the use of museum and other archival collections for deep-sea research is still rare although it offers an immense potential for different kinds of investigations.

Generally, in order to conduct molecular research on deep-sea organisms, there are two ways to obtain samples for a study:

**(a)** use samples from existing archival collections, and/or

**(b)** collect samples in the field.

Special requirements and pre-conditions apply for either way of securing samples for such projects. Because of the extreme rarity of having frozen archival collections of a range of deep-sea organisms available for molecular investigations (Schander and Halanych 2003), it is more feasible to conduct research on preserved specimens from museum and other archival collections. The majority of these collections contain specimens that are fixed in formalin (Schander and Willassen 2005; Skage and Schander 2007). This means investment of substantial time and laboratory work in developing molecular approaches for obtaining molecular information from such specimens (Schander and Willassen 2005; Tang 2006). On the other hand, collecting samples in the field requires research teams that possess specially equipped boats with a crew trained for deep-sea fishing and collecting samples for molecular research (Schander and Halanych 2003; Gage 2004). Deep-sea equipment is very expensive and not many research/academic institutions around the world have the necessary equipment and trained teams. Thus, in order to expand research and broaden knowledge about deep-sea organisms, it is more advisable to use already existing and well-documented archival collections with reliable identification of species, than to wait for the rare

opportunity to go on a deep-sea exploration boat as a part of not frequently organised and very expensive expedition.

A serious threat to achieving comprehensive knowledge about the ocean is also the human destructive activity by using the sea as a resource for humans (exploitation of living and non-living resources) and as a dumping place for all kinds of wastes, as well as the introduction of alien and non-native species (see Probert *et al.* 1997; Haedrich and Barnes 1997; Jennings and Kaiser 1998; Creasey and Rogers 1999; Koslow *et al.* 2000; Beznosov and Suzdaleva 2001; Monteiro *et al.* 2001; Thrush and Dayton 2002; Allain *et al.* 2003; Morley *et al.* 2004). These have direct impact on the life in the sea locally, but extensive and continuous anthropogenetic influence of this sort might severely affect the existence and population structure of species in the ocean, including the deep sea (see Grassle 1977; Mortensen *et al.* 1995; Merrett and Haedrich 1997; Rogers 1999; Gordon 2001; Thiel 2003; Kropp 2004). Increasing human impacts on marine habitats might have a severe effect on deep-sea organisms with the consequence that many of them become endangered or even extinct before we know about them (Haedrich *et al.* 2001; Key 2002). The bathyal zone of the deep sea and its species diversity might be also impacted by global climatic and oceanographic changes (Cronin and Raymo 1997; Rex 1997; Key 2002; Genner *et al.* 2004; Nunes and Norris 2006), although the exact impact and all possible consequences for deep-sea communities are still unknown (Glover and Smith 2003; Schiermeier 2006).

## 1.8. *Introduction to Nezumia and its relevance in marine biology*

The *Nezumia* is a species-rich genus (including about 60 species; www.fishbase.org) and an important constituent of the deep sea (Coggan *et al.* 1999). However, species of this genus are genetically uncharacterised and only basic morphological, ecological and population structure data are available (Marshall and Iwamoto 1973; Cohen *et al.* 1990; Coggan *et al.* 1999; Ross *et al.* 2001; Madurell *et al.* 2004; Merrett, personal communication). Little is known about their reproduction and early life history, the migratory capability of adults and dispersal capacities of larvae and young fish (Merrett 1986, 1989; Merrett and Barnes 1996; Merrett 2003). Patterns of evolution and phylogenetic relationships within this genus and macrourid fish are in general also unclear (Iwamoto and Merrett 1997, personal communication; Fig. 3). It is surprising that such a big constituent of the deep sea has so little information and no genetic information available.

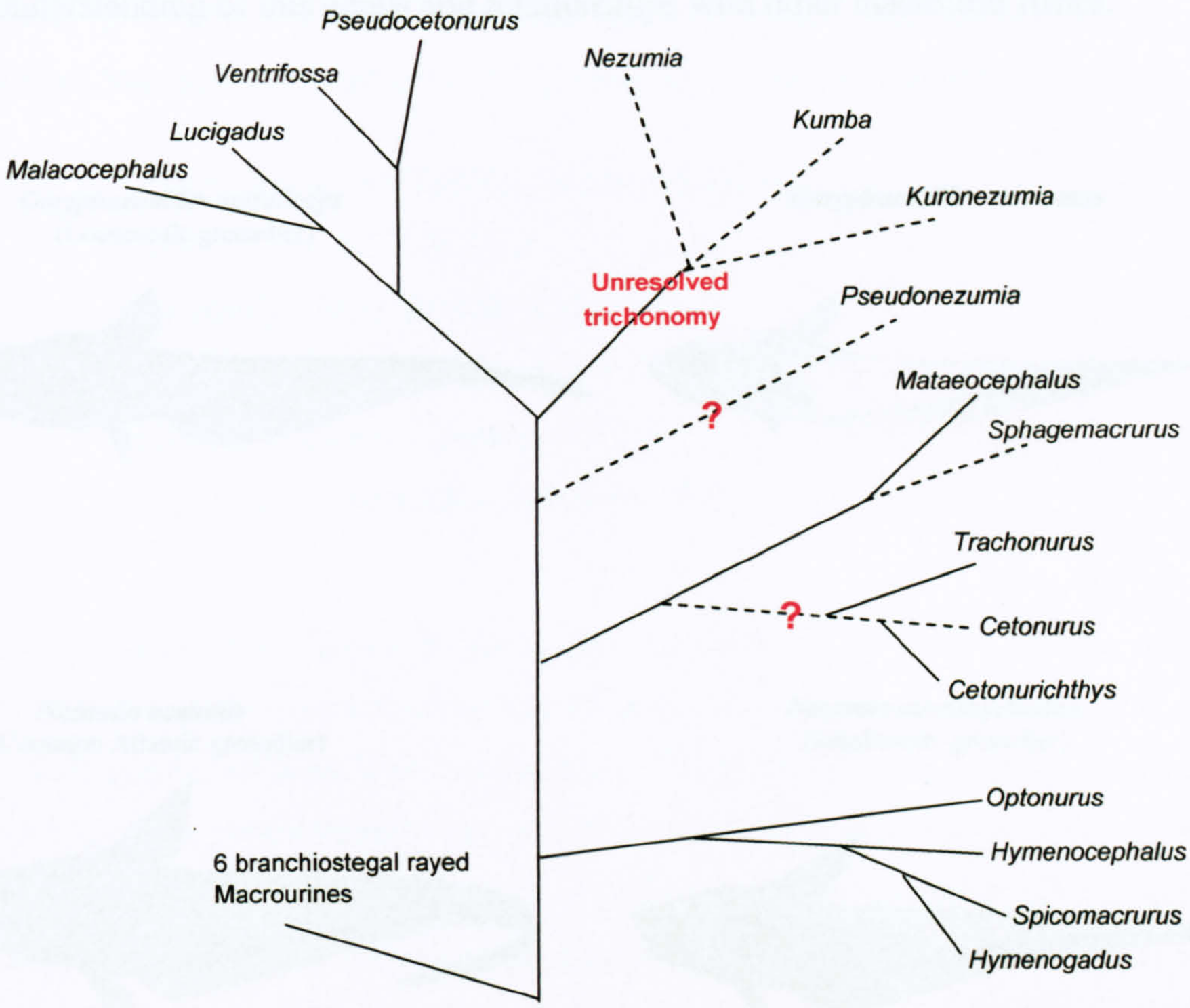
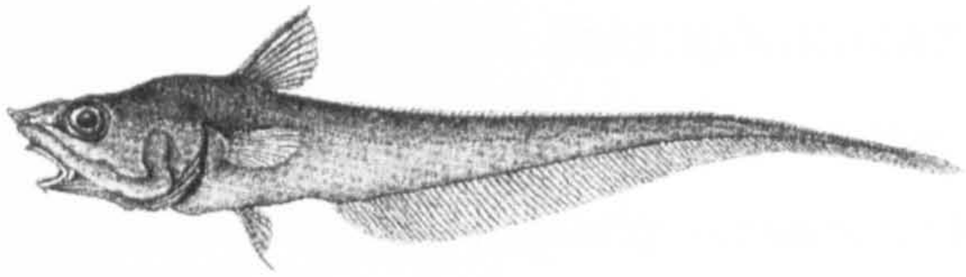



**Fig. 3** – Iwamoto`s (data) and Merrett`s (tree) view on phylogenetic relationships of macrourid fish based on morphological characters (number of branchiostegal ray). The particular groups, i.e. their interrelationships are currently not be able to be completely resolved (Iwamoto and Merrett 1997, personal communication). Regarding phylogenetic relationship of the genus *Nezumia,* it probably belongs in a clade that includes *Kumba* and *Kuronezumia,* which in turn is most closely related to the clade that includes *Malacocephalus, Lucigadus* and *Ventrifossa;* and *Pseudonezumia* may be a sister group of all these (Iwamoto and Merrett 1997, personal communication).

The genus *Nezumia* is an example of taxon that comes from a remote area (the deep sea) and is available mostly as archival specimens. But, *Nezumia* specimens with full and reliable identification at the level of species (or even genus) are rare in collections (preserved and frozen) because specimens in collections are often with a partial and/or uncertain identification if experts for these groups of fishes have not been involved in their identification (see Iwamoto *et al.* 1999). This causes problems for their broader usage for different biological studies (including molecular investigations). For example, for non-experts on macrourid fish, the relatively large genus *Coryphaenoides* has a morphological appearance very similar to *Nezumia* fishes (Figure 4), especially as alevins (young fishes) or

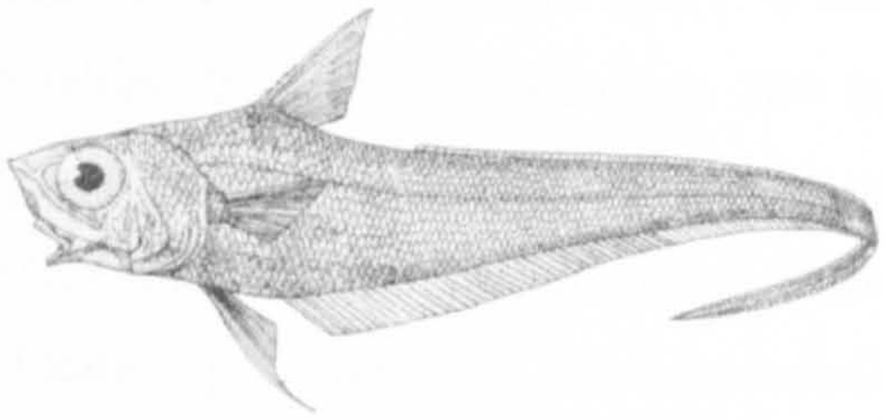
larvae. All these indicate the urgent need for developing appropriate molecular markers for their DNA identifications and molecular investigations (accumulating sequence data in particular) in order to include these species in broader biological studies that will help in a better understanding of this genus and relationships with other macrourid fishes.

***Coryphaenoides anguliceps***
(Loosescale grenadier)

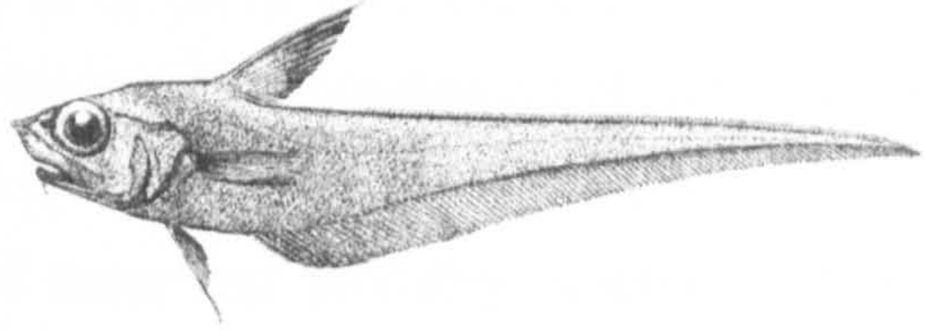


***Coryphaenoides ariommus***

***Nezumia aequalis***
(Common Atlantic grenadier)



***Nezumia micronychodon***
(Smalltooth grenadier)

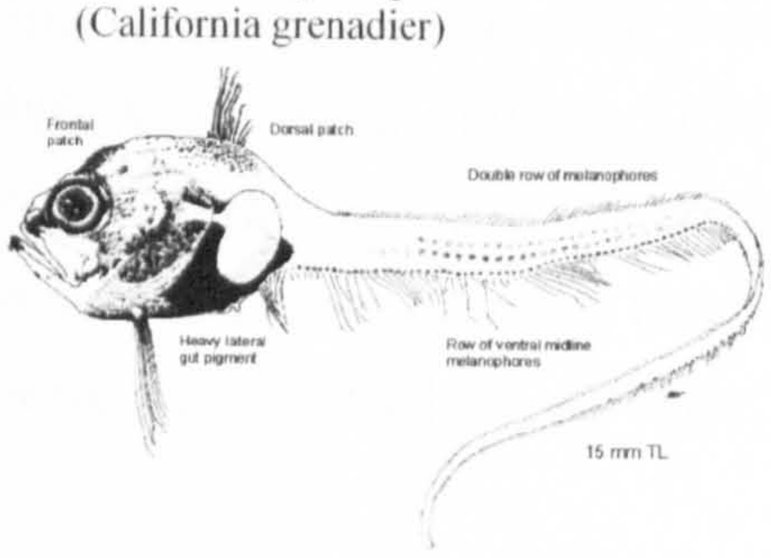


(from: www.fishbase.org)

---

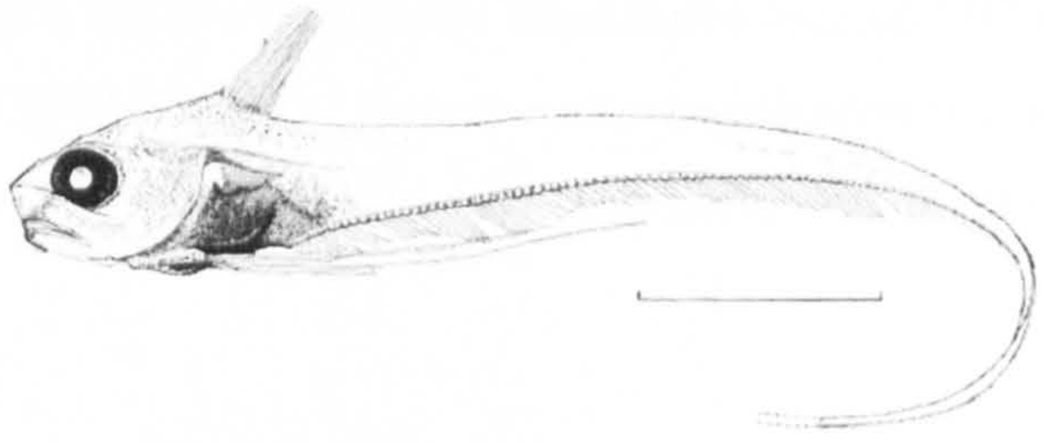
## Early Life History Photos

***Nezumia stelgidolepis***
(California grenadier)




(from: www.fishbase.org)

***Nezumia aequalis***
(Common Atlantic grenadier)



(from: Merrett 1989)

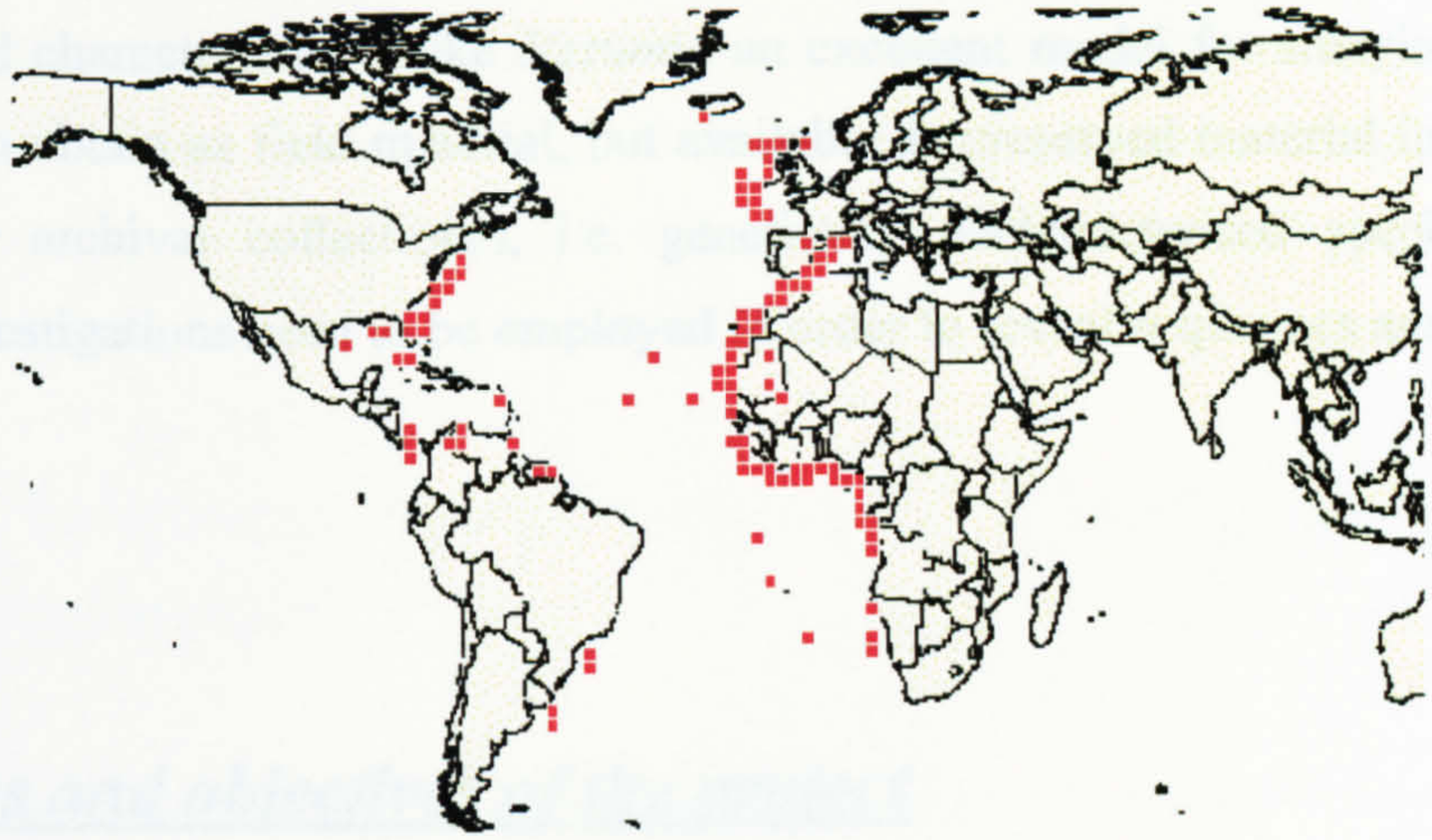
**Fig. 4.** – Some rattail (grenadier) fish as adults and in early life history (larvae or alevins)

*Nezumia* is not important from an economic point of view (in fisheries), but it is important from a fundamental biological aspect as being a significant constituent of the deep sea (Coggan *et al.* 1999). The genus *Nezumia* and two species investigated in this project, according to Marshall and Iwamoto (1973), have the following taxonomic order:

Order: ANACANTHINI (GADIFORMES)
Family: MACROURIDAE
Subfamily: MACROURINAE
Genus: *Nezumia* (Jordan, 1904)
Species: *aequalis* (Gunther, 1878)
*micronychodon* Iwamoto, 1970

The geographical distribution of *Nezumia aequalis* (Figure 5 (a)) is as follows: widespread in North Atlantic, from Faroe Bank to northern Angola and Mediterranean Sea in the east, and Davis Straits to northern Brazil in the West (Cohen *et al.* 1990). This is a benthopelagic fish which lives in approximately 200 to 1000 m of depth with size to about 30 cm total length (Cohen *et al.* 1990). Cohen *et al.* (1990) stated: "Populations of this widespread North Atlantic species have been distinguished by Marshall and Iwamoto (1973). The Gulf of Guinea population is the most distinct of these and probably deserves subspecific recognition at the least. The population in the Gulf of Mexico, although less well defined, also deserves consideration at the subspecific level."

Another species that has been used for this study is *Nezumia micronychodon* Iwamoto 1970. This has a total length of the body up to 34 cm and a geographical distribution around West Africa (Figure 5 (b)), i.e. from Western Sahara (about 24°N) south to Angola (Cohen *et al.* 1990). Adults live in 360 m to 1620 m depth, but are in greatest abundance around 500 m to 600 m (Cohen *et al.* 1990). Cohen *et al.* (1990) remark: "*Nezumia micronychodon* is readily distinguished from all other members of the genus by its high gill-raker count and the relatively unrestricted gill openings."



**(a)** ***Nezumia aequalis*** geographical distribution (from: www.fishbase.org)

**(b)** ***Nezumia micronychodon*** geographical distribution (from: www.fishbase.org)

**Figure 5** – Geographical distributions of *Nezumia aequalis* (widespread species) and *N. micronychodon* (regional distribution) (from: www.fishbase.org)

As already known, accumulation of genetic changes within and among populations may lead to the formation of new subspecies or even new species. These genetic changes (genotypes) might not be clearly visible in the changing of phenotypes, so measurement of morphological variation on its own is not always the best criterion to identify species and the diversity within and between populations (Liu and Cordes 2004; Schander and Willassen 2005). Cohen *et al.* (1990) already remarked on the necessity for additional investigations on *Nezumia aequalis* populations at the subspecific level, because the populations in the Gulf of Guinea and the Gulf of Mexico express more distinctive characteristics than other populations of this widespread species.

All mentioned characteristics make *Nezumia* an excellent model for studying species that are difficult to obtain as field material, but available as preserved material in the museums (and/or other archival collections), i.e. genetically uncharacterised species for which molecular investigations need to be employed in order to reveal sequences and basic genetic structure.

## *1.9. Aims and objectives of the project*

This project involved molecular work on preserved museum specimens of deep-sea fish (*Nezumia aequalis* and *N. micronychodon*).

The **main objectives** of this study are:

***(a) to develop appropriate molecular approaches for studying formalin-fixed, Steedman`s-preserved fish specimens using the collection from the Natural History Museum, London, and***

***(b) to provide the first molecular data on two deep-sea fish species (Nezumia aequalis* and *N. micronychodon*).**

Other aims of the project were:

- To test whether DNA could be recovered from formalin-fixed museum samples, and whether it is in a condition suitable for PCR amplification and sequencing. Results obtained in this study should provide information on the feasibility of conducting a research project solely on museum specimens in order to evaluate and demonstrate the possibility to accumulate reliable molecular data.

- To test the efficiency of different DNA extraction protocols using particular chemicals or reagents, types of tissue, and testing tissue pre-washing and drying regimes prior to DNA extractions in order to obtain DNA extracts of good yield and quality. The DNA extracts were tested in association with PCR effectiveness.

- To test different molecular techniques and marker systems for investigations of nuclear and mitochondrial genome using PCR technology, Southern blotting hybridisation, cloning and sequencing in order to estimate the reliability of the particular techniques and methodologies.

- To test the efficiency of DNA extractions/RAPD amplifications by using RAPD methodology approach.

- To test the reliability and reproducibility of RAPD data extracted from formalin-fixed specimens of fish.

- To investigate the possibility of identifying species-specific, or genus-specific markers.

- To design specific primers that will allow sequence comparison between *Nezumia aequalis* and *N. micronychodon*, and other species.

- To validate data obtained in different laboratories and on differently preserved specimens (DMSO and ethanol) that come from different cruises, locations and collections.

- To interpret and make valid comparisons of the different strands of the data.

- To establish an initial molecular-genetic dataset for *Nezumia aequalis* and *N. micronychodon* that will be useful for the further studies of these species.

One of the molecular objectives in this thesis is to explore the potential and the reliability of RAPD-PCR technology and RAPD markers in general, knowing that this methodology is the subject of over a decade of debate on its reliability and reproducibility. The special challenge was to investigate these issues applying RAPD-PCR technology in the research that was solely based on studying museum formalin-fixed specimens of fish.

This study tried to put together and investigate two important aspects: (1) the use of preserved fish specimens in molecular investigations, and (2) to contribute in the genetic characterisation of deep-sea fish (genus *Nezumia*) toward our better understanding of the genetics of the investigated species.

# Chapter 2. GENERAL MATERIALS AND METHODS

## *2.1. Specimen collection, identification and preservation*

The formalin-fixed specimens of *Nezumia aequalis* and *N. micronychodon* (Family: Macrouridae) were used for the main study, but fresh/frozen, ethanol and DMSO preserved specimens of these and other fish species were included as controls for the validation of results. Ten differently preserved fish specimens of Atlantic mackerel (*Scomber scombrus*, Linnaeus, 1758) were used as supplementary material to investigate DNA extraction protocols and different preservation regimes.

### 2.1.1. Main study

Samples of formalin-fixed deep-sea fishes of *Nezumia aequalis* (Gunther, 1878) and *Nezumia micronychodon* Iwamoto, 1970 (Fig. 6) for this study were obtained from the well-documented collection of the Natural History Museum, London, U.K (NHM). *Nezumia aequalis* samples were collected in July 1979 (51°01.7`N, 14°12.1`W) from 700 m – Station 50607#01 (Dr Nigel Merrett from the Natural History Museum, personal communication); Registration no.: BMNH 1998.8.9.378. *Nezumia micronychodon* samples were collected in July 1987 (20°41.1`N, 17°56.9`W - North-West Africa water) from 840-930 m – Station 11542#01 (Dr Nigel Merrett, personal communication); Registration no.: BMNH 2002.3.1.742-743. Samples were collected by trawling. Taxonomic identification based on morphological characters was carried out by Dr Nigel Merrett.

Specimens of both species were fixed in unbuffered 10 % sea-water formalin for 24 hours, and then transferred to Steedman's preservative that contained 150 ml formalin, 1500 ml propylene glycol, 75 ml propylene phenoxetol and 13275 ml distilled water (Dr Nigel Merrett, personal communication). Samples of *N. aequalis* were stored in this solution at room temperature for almost 20 years, and *Nezumia micronychodon* for about 10 years.

Investigated fish specimens ranged in length from 95 mm to 330 mm for *N. aequalis* (for 19 fish specimens), and from 160 mm to 300 mm for *N. micronychodon* (for 12 fish specimens).


**Figure 6** – Formalin-fixed, Steedman's-preserved museum specimens of the two species used in this study

### 2.1.2. Other specimens

A small DMSO preserved muscle tissue sample of *Nezumia* cf. *aequalis* caught off Southeast Greenland in 1999 (62°88`N and 40°34`W from a depth of about 900 m) was obtained from the collection of the Greenland Institute of Natural Resources through the Natural History Museum, London, U.K.

A two and half year-old ethanol preserved (70% ethanol) muscle tissue sample of *Nezumia aequalis* was obtained from Dr Martin Collins from the British Antarctic Survey, Cambridge, UK through the Natural History Museum, London (fish samples were caught

49°43`N and 11°46`W from depth of about 1050 m), and 1 – 1.5 year-old ethanol preserved tissues of *Nezumia aequalis* and *Nezumia* cf. *micronychodon* obtained from Dr Reinhold Hanel from the Institute fur Meereskunde, Kiel, Germany (caught off the South Africa coast at a depth of 750 m for *Nezumia aequalis*, and 320-385 m for *Nezumia* cf. *micronychodon*; samples obtained through the Natural History Museum, London).

Fresh/frozen cod (*Gadus* sp.; Family: Gadidae) and rainbow trout (*Oncorhynchus mykiss* (Walbaum, 1792); Family: Salmonidae) were included in these experiments as a control material (unpreserved tissue of fish) for testing extractions of DNA and PCR amplifications. Ethanol preserved tissue of *Coryphaenoides armatus* (Hector, 1875) and *Coryphaenoides rupestris* Gunnerus, 1765 (obtained through Dr Alex Rogers, from collections of the Oceanography Centre, Southampton University), as well as DNA of these two species (phenol extraction; provided by Dr Alex Rogers, Molecular Ecology Group, Oceanography Center, Southampton University) were used in some experiments as a control. Samples of *Coryphaenoides armatus* had been collected in April 1999 from the Porcupine Abyssal Plain in North-East Atlantic Ocean by trawl and immediately fixed and preserved in 95 % ethanol (Dr Alex Rogers, personal communication).

### *2.1.2.1. Supplementary specimens*

Four-year-old collections of ten differently preserved mackerel (*Scomber scombrus*, Linnaeus, 1758) specimens from the Natural History Museum, London, U.K, were used as supplementary specimens to the main study (Fig. 7), for testing different preservation regimes in relation to DNA extractions and PCR amplifications, as well as for validation of some results from the main study. These fish were collected exclusively for molecular studies respecting all necessary rules of collecting specimens for this kind of study: the fish were preserved immediately upon collection, the specimens were handled with gloves and all data about catching and preserving fish specimens are recorded in detail.

- ***100% ethanol,*** (Ref. no. 2003.9.30.7)
- ***70% ethanol,*** (Ref. no. 2003.9.30.6)
- ***95-100% industrial methylated spirits (IMS),*** (Ref. no. 2003.9.30.10)
- ***70% IMS,*** (Ref. no. 2003.9.30.4)
- ***50% isopropanol,*** (Ref. no. 2003.9.30.3)
- ***Steedman`s solution,*** (Ref. no. 2003.9.30.5)
- ***3% unbuffered formalin,*** (Ref. no. 2003.9.2)
- ***3% buffered formalin (with 4% sodium tetraborate),*** (Ref. no. 2003.9.30.8)
- ***10% unbuffered formalin,*** (Ref. no. 2003.9.30.1)
- ***10% buffered formalin (with 4% sodium tetraborate).*** (Ref. no. 2003.9.30.9)

**Figure 7** – Ten differently preserved specimens of mackerel (4-year-old collections); supplementary specimens

## 2.2. ***DNA extraction and molecular investigations***

A range of DNA extraction protocols, pre-extraction treatments of fixed tissue, molecular methods and techniques were applied in this study in order to test their applicability and usefulness in molecular investigations on preserved specimens, that is, in order to optimise and develop an adequate DNA extraction protocol and molecular approaches for studying formalin-fixed, Steedman's-preserved specimens of fish. Care was taken at all stages to avoid contamination by human and other external DNA present in a molecular laboratory.

### 2.2.1. Precautions undertaken in order to avoid contaminations with exterior sources of DNA and cross contaminations

Disposable sterile plastic ware and pipette tips were used throughout the experiments, along with sterile mortar and pestle, aliquoted sterile reagents and solutions dedicated solely for the work on this project and DNA extracted from museum specimens. The bench surfaces, outsides of boxes with pipette tips, pipettors, outsides of bottles of reagents, or anything that was used in connection with DNA extraction of these samples but not disposable, were regularly cleaned by distilled water and 2% hycolin, or 5-10% bleach. Special care was taken so that none of these "aggressive" chemicals (hycolin and bleach) could get into direct contact with reagents and tissue samples used for DNA extraction.

Blank control extractions (extractions performed with $sdH_2O$ instead of using tissue) were carried out in parallel with the DNA extractions of preserved tissue in order to monitor contamination which might come from reagents, laboratory glassware and plastic ware. Positive control extractions (containing frozen tissue of cod, fresh/frozen tissue of rainbow trout, or ethanol preserved *Coryphaenoides*) were never carried out simultaneously with DNA extractions from formalin preserved *Nezumia* specimens in order to prevent possible cross contamination between positive controls and experimental, preserved samples. PCR blanks – negative controls of PCR reactions (using sterile water as a template instead of DNA) were performed with each PCR experiment.

All reagents used for DNA extraction of positive control tissues were separated from the reagents used solely for DNA extractions of *Nezumia* formalin-fixed tissues in order to prevent possible cross contamination between positive and experimental samples through the reagents. Also, lab coats used in the procedures for DNA extractions of positive and preserved tissues were kept separate. Any lab coat used for setting up PCR reactions was always kept away from any lab coat in the laboratory by keeping it in a plastic bag in a cupboard dedicated only for keeping pipettors and plastic ware used for PCRs. All these precautions were necessary to prevent cross contamination and contamination with exogenous sources of DNA.

PCR experiments were carried out under strictly controlled conditions in a laboratory in order to prevent accidental contamination with foreign, extrinsic DNA and "PCR product

carry-over". Pipettors used for PCRs were separated from pipettors used for other associated procedures (DNA extractions, cloning, etc.), and pipettors were not exchanged between researchers in the Laboratory. For PCR set up, only pre-sterilised filter tips (ordinary aerosol [Sigma], or Rainin shaft-guard aerosol pipette tips, [Anachem]) and γ, UV pre-sterilised PCR tubes (Advanced Biotechnologies) were used. All these were kept and stored in a separate cupboard from those used in other molecular applications. Boxes with gloves that were used for PCR experiments were stored in the "PCR cupboard", and if powdered gloves were used, the gloves were washed prior to any PCR work. All racks and containers for ice were separate from others, kept in the "PCR cupboard" and used only for PCR experiments, washed each time after use and UV sterilised (5-10 minutes on transilluminator) before use. Before PCR experiments, or any work related to the PCR, the laboratory bench would be wiped with 5-10% bleach and/or 2-5% hycolin solution.

Special, pre-sterilised water (18 Megohm Water, Sigma), divided into aliquots, was additionally UV-treated on UV-transilluminator for 5-10 min and used for all PCR experiments and steps related to PCR work (diluting reagents and preparing stock solutions for PCRs) in order to prevent contamination through the water. All reagents used for PCRs were aliquoted and, most of them, after being used once for PCR set up were discarded (magnesium, buffer, dNTPs, water, and so on). Only working stock solutions of primers and *Taq* were aliquoted in a way that could be used more than once (but not for more than three PCR set ups). *Taq* PCR kits were not exchanged between researchers in the laboratory.

### 2.2.2. DNA extraction protocols and pre-extraction treatments of preserved tissues

During the research, various DNA extraction protocols (commercial kits and protocols developed by other researchers) and different pre-washing/drying regimes for tissue samples prior to DNA extraction were tested.

DNA extraction protocols used in this study and a newly developed protocol during this study are described below. For all protocols and in all laboratories, an Eppendorf microcentrifuge (model 5415C) was used.

## Protocol (A)

**Guanidinium-based protocol.** A newly developed protocol (A) was developed by combining and modifying two DNA extraction protocols: one of Hammond *et al.* (1996) - protocol (B), and another of Shedlock *et al.* (1997) – protocol (C). The protocol developed in this study is outlined in detail below:

1. Tissue was cut into a small piece (ca. 0.5 $cm^3$) with a sterile razor blade, and then interchangeably washed in distilled water and in freshly made filter sterilised pre-washing 1xGTE buffer (100 mM glycine, 10 mM Tris-HCl pH 8.0, 1 mM EDTA; Shedlock *et al.* 1997) for 72 hours at room temperature (using a suspension mixer for gentle shaking), frequently changing water and pre-washing buffer during tissue washing.

2. The tissue was air-dried for ~30 min at room temperature. Dried tissue was placed in a pestle and mortar with alumina powder for grinding. Traces of integument were completely removed with sterile forceps. Solid $CO_2$ was added. When the sample had been ground to a fine powder, a small amount of alumina powder was again added for final grinding. When the solid $CO_2$ had sublimed, 500 µl of extraction (incubation; digestion) buffer (1% SDS, 25 mM Tris-HCl pH 7.5, 100 mM EDTA; Shedlock *et al.* 1997) was added and grinding was continued for a few more minutes.

3. The mixture was transferred to a sterile 1.5 ml Eppendorf tube. 100 µl of proteinase K (10 mg/ml; Sigma, St. Louis, MO, USA) and 20 µl of 1 M dithiothreitol (DTT; NBL Gene Sciences, Northumberland, UK) was added. The tube with this mixture was placed in a water bath overnight at 55°C.

4. An additional 50 µl of Proteinase K (10 mg/ml; [Sigma] or [Gibco BRL, Life Technologies]) and 10 µl of DNase-free RNase A (10 mg/ml; Sigma) was added next day; that is, after at least 10-15 hours of incubation from the previous step. The digestion was continued further for about 10 hours, i.e. until the tissue was completely digested (pieces of tissue were not visible).

5. After digestion was completed, 500 µl of the GES reagent (0.5 M guanidinium

thiocyanate and 0.1 M EDTA; Hammond *et al.* 1996) was added at room temperature. It was mixed for few minutes by gently inverting and flicking the tube.

6. Then, 250 µl of 7.5 M ice-cold ammonium acetate was added, gently mixed by inverting the tube, and then immediately transferred to ice to stand for 10 min.

7. Chloroform : Pentanol, 24 : 1 (500 µl) was added at room temperature and gently mixed with the samples by inverting the tube. Then, centrifugation was applied at 13,000 rpm for 10 min. The upper, aqueous phase from the tube was transferred to a new sterile 1.5 ml Eppendorf tube without disturbing the precipitated material at the interface.

8. For DNA precipitation, 0.54 vol. of 100% isopropanol was added and left overnight at -20°C (or at 4°C). Then, the DNA precipitate was spun for 30 min in a microcentrifuge at 13,000 rpm.

9. The supernatant (SN) was removed and the pellet was rinsed three to five times with 70% ethanol. After ethanol was carefully discarded, the DNA pellet was air dried at room temperature for 10-15 min (the tube was inverted on clean, UV-treated absorbent paper). The DNA pellet was re-suspended in 40-45 µl of 1xTE buffer (pH 7.6).

The protocol described above is suited to obtain a volume of 40-45 µl of DNA suspension (extract) per one DNA extraction procedure, but in order to produce a larger amount of DNA suspension (extract) per one DNA extraction procedure (that would allow a higher number of PCR amplifications with the same DNA extract and eliminate possible differences between DNA extracts obtained with the same protocol, but in separate DNA extraction procedures), some adjustments in the described procedure were required:

- A bigger piece of preserved tissue was used (usually 1-3 $cm^3$). The incubation (digestion) buffer (Shedlock *et al.* 1997) was gradually added into a pestle with ground preserved tissue (continuing to grind tissue using a mortar) added until a relatively thick mixture was obtained. The mixture was transferred (using wide-open, 1 ml pipette tips) into a series of 1.5 ml Eppendorf tubes. The volume of mixture was about 500-700 µl in each tube.

The rest of the procedure was carried out according to the above-described protocol, but with small adaptations. At step **7**, aqueous phases from two or more tubes were collected in fresh tubes filling them up to 1000 µl. Then, 0.54 vol. of 100% isopropanol was added into each tube as described in the above protocol (Step **8**). After completing step **9**, DNA

suspensions from each tube were put together into one 1.5 ml Eppendorf tube, mixed and quantified. Afterwards, the DNA suspension was immediately aliquoted (usually 50 µl of DNA suspension in each tube). All aliquots made in this way (in one extraction procedure and from one piece of tissue) were considered as one extraction.

***<u>Protocols used for developing a new protocol (A)</u> – the protocols (B) and (C)***

## Protocol (B)

**Guanidinium-based protocol.** This protocol, described by Hammond *et al.* (1996), is based on the fine grinding of the tissue in a mortar using alumina powder and solid $CO_2$, and then DNA extraction by guanidinium thiocyanate with the aid of ammonium acetate. The original protocol (Hammond *et al.* 1996) is briefly outlined below:

1. Skeletal muscle (2-3 mm cube) from a preserved fish was washed in 95% ethanol, and then the tissue was rinsed and soaked in sterile water.

2. The tissue was ground in a pestle and mortar with solid $CO_2$ and a small amount of alumina added for final grinding.

3. After the solid $CO_2$ had sublimed, but before the sample had thawed, 500 µl of extraction solution – the GES reagent (guanidinium thiocyanate 0.5 M; EDTA 0.1 M) was added.

4. The sample suspension was transferred to a sterile 1.5-ml Eppendorf tube. Ice-cold ammonium acetate (250 µl; 7.5 M) was added and mixed with the suspension.

5. Chloroform reagent (500 µl; chloroform:pentanol, 24:1) was then added and mixed with the sample.

6. Isopropanol (0.54 vol) was used for DNA precipitation and 70% ethanol for the pellet washing.

For this study with *Nezumia*, the Hammond *et al.* (1996) protocol was used as described, or with modifications such as:

- For pre-washing tissue samples, only sterile distilled water was used without prior ethanol;
- Alumina powder (used for grinding tissue in a pestle and mortar) was added before and after adding dry ice;
- The Promega Wizard DNA Clean-Up System (Promega, Madison, USA) was tested instead of using isopropanol for DNA precipitation;
- The DNA pellet was washed five times instead of three times with 70% ethanol.

## Protocol (C)

**Phenol-based protocol.** Another protocol tested in this study, and partially used for developing the protocol (A), is the one described by Shedlock *et al.* (1997):

1. Small pieces of formalin-fixed muscle tissue (ca. 0.5 $cm^3$) were dissected with sterile razor blades. Traces of integument were completely removed, and then the pieces were washed for three successive 24-hour periods (rotary shaker at room temperature) in a fresh solution of 10 ml of 1x GTE pre-washing buffer (100 mM glycine, 10 mM Tris-HCl, pH 8.0, 1 mM EDTA) to act as a binding agent for excess formalin.

2. Tissue was air-dried and completely digested in 500 µl of extraction buffer (1 % sodium dodecyl sulphate [SDS], 25 mM Tris-HCl, pH 7.5, 100 mM EDTA) at 65°C for 24 h. Then, 20 µl of 1 M dithiothreitol (DTT) and 100 µl of proteinase K (10 mg/ml) were added at the beginning of the digestion. An additional 50 µl of proteinase K (10 mg/ml) and 10 µl of DNase-free RNase (10 mg/ml) were added after the first 10 h of digestion.

3. Phenol/chloroform extractions were performed as follows:
   (a) completed digestions were extracted in 500 µl of equilibrated phenol. Supernatants were saved and extracted two more times in 500 µl of equilibrated phenol;
   (b) supernatants were extracted twice with 500 µl of 25:24:1 solution of phenol : chloroform : isoamyl alcohol; and
   (c) supernatants were extracted twice with 500 µl of 24:1 solution of chloroform : isoamyl alcohol.

4. DNA in supernatants was precipitated by adding 2.5 vol of cold absolute ethanol (stored at –80°C) and samples were immediately placed at -20°C for 24 h.

5. DNA precipitates were spun for 30 min in a microcentrifuge at 13,000 rpm.

6. Absolute ethanol was removed. Pellets were rinsed twice with 50 µl of 70 % ethanol and thoroughly air-dried.

7. Purified DNA was re-suspended in 40 µl of 1 x TE (pH 8.0).

Major modifications applied to this protocol:

- For step **1**: different lengths of tissue pre-washing were tested (from 30 min to up to 72 hours). Tissue pre-washing with only sterile water was tested instead of using 1xGTE buffer, or interchangeable washing with sterile water and 1xGTE buffer.

- Step **2**: after drying, the tissue was ground in a pestle and mortar with solid $CO_2$ and a small amount of alumina powder;

- Step **3** (a) was modified so that the phenol was used only once instead of three times, or complete step **3** (a) was omitted,

- Step **3** (b) was modified in a way that phenol : chloroform : isoamyl alcohol was used only once instead of twice, or complete step **3** (b) was omitted,

- Step **3** (c) was modified in a way that the chloroform : isoamyl alcohol was used only once instead of twice,

- For the step **3**, phenol was not used at all, i.e. steps **3** (a) and **3** (b) were omitted completely and only step **3** (c) was applied;

or

Complete step **3** was omitted; that is, GES reagent from protocol (B) and chloroform reagent, or the Promega Wizard DNA Clean-Up System Kit (Promega, Madison, USA) was used instead.

- Step **4**: DNA precipitation was tested with 2.5 vol of absolute ethanol + 1/10 vol of 2 M NaCl, or with 0.54 vol of isopropanol.

- Step **6**: Extended pellet rinsing, 3, 4 or 5 times with 50 µl of 70 % ethanol. Different amounts of 70% ethanol (50 -750 µl) were applied for testing for the pellet-washing step.

***Commercially available reagents (kits) for DNA extraction tested at the NESCOT Laboratory – protocols (D) and (F)***

## Protocol (D)

**Genomic DNA Purification Reagent – Genosys DNA Isolator** (Genosys Biotechnologies, Inc., Cambridge, UK) was tested according to the manufacturer's protocol (unmodified) and with protocol modifications. The basic steps of this protocol (according to the manufacturer's protocol) are:

1. HOMOGENIZATION: Lyse sample by adding 1 ml of DNA Isolator reagent (per 50 - 100 mg) and incubate homogenate for 5 minutes at room temperature.

2. PHASE SEPARATION: Add 0.2 ml chloroform, mix and incubate for 5 minutes at 4°C. Centrifuge for 2-5 min at 5,000-12,000 rpm.

3. DNA PRECIPITATION: Transfer aqueous phase into a new tube and add 1-2 volumes of isopropanol (0.5-1.0 ml). Mix gently by inversion and incubate for 10 minutes at 4°C.

4. DNA WASH: Centrifuge for 5 minutes at 5000 rpm. Remove supernatant and wash DNA pellet with 1.0 - 1.5 ml 70 % ethanol. Repeat ethanol washes two more times.

5. RESUSPENDING DNA: Air dry the DNA pellet for 5 - 10 minutes, then dissolve in 50-200 µl of sterile water, or in 1 x TE buffer.

Applied modifications on the protocol were:

- Prior to the DNA extraction the pieces of formalin-preserved tissue were washed in sterile water for 24 - 36 hours at room temperature (frequently changing the water).
- After adding 1 ml of "DNA isolator" reagent, tissue was placed in tubes with glass beads to shake at medium speed for 5 - 10 s in the Ribolyser machine (Hybaid). Afterwards, tissue was kept in "DNA isolator" for up to 1 hour at room temperature.

or

The tube containing 1 ml of "DNA isolator" reagent and macerated preserved tissue were placed on to a suspension mixer at room temperature for up to 1 hour (for a gentle agitation by inverting the tube at low speed).

- DNA precipitation with 1 ml of isopropanol was extended from 10 min to overnight (in a fridge at 4°C)
- Centrifugation speed was increased from 5,000 rpm to 12,000 rpm at step 4.

- DNA pellet was air dried at room temperature for up to 30 min, or at 37°C for 15 minutes.
- The pellet was re-suspended in 50 µl of sterile water.

## Protocol (F)

**The Promega Wizard Genomic DNA Purification Kit** (Promega, Madison, WI, USA) was another commercially available kit tested for DNA extractions from formalin fixed specimens. The protocol for animal tissue (mouse tail) with small modifications to the original protocol was applied. Modifications were related only to washing tissue with sterile water and grinding a tissue sample in a pestle and mortar.

This protocol was tested only once with one *Nezumia aequalis*-preserved tissue piece. A tissue sample for this extraction was previously washed in sterile water and stored for a few months at -64°C. To extract DNA, a tissue sample was thawed and briefly air-dried at room temperature (for a few minutes), and then placed in a pestle and mortar with alumina powder for grinding. Traces of integument were removed with sterile forceps. Solid $CO_2$ was added and grinding continued. When the sample had been ground to a fine powder, a small amount of alumina powder was again added for a final grinding. Powdered tissue sample was transferred into 1.5 ml microcentrifuge tube with prepared ice-cold lysis solution (mixture of 120 µl of 0.5 M EDTA (pH 8) + 500 µl of Nuclei Lysis Solution from the Kit). The extraction then followed the manufacturer's instructions.

***<u>The protocols and DNA extraction kits tested in the Molecular Ecology Group at the Oceanography Centre of the Southampton University</u> - protocols: ($G_1$), ($G_2$), and ($G_3$)***

DNA extraction methods which have been used regularly by the Molecular Ecology Group at the Oceanography Centre, University of Southampton, for frozen and short period (few months) ethanol preserved samples of deep-sea organisms (Rogers 1999; personal communication) were tested on formalin-fixed samples of *Nezumia*. The methods tested are described below:

## Protocol ($G_1$)

**Organic (phenol/chloroform/isoamyl alcohol; PCI)** extraction of DNA was tested with and without modifications.

<u>The original Southampton's PCI protocol</u>:

1. Aliquot 600 µl of extraction buffer (40 µl of 1M Tris-HCl (pH 8.0) + 20 µl of 25% SDS + 350 µl $H_2O$) into 1.5 ml an Eppendorf tube and add a small piece of tissue (about 0.5 $cm^3$) and macerate gently. Add 10 µl Proteinase K Solution of 15.6 mg/ml (Boehringer Mannheim) (or 5 µl of 20 mg/ml proteinase K). Incubate for 2-3 hrs at 55°C inverting tubes every 30 minutes.

2. In a fume hood add 300 µl of phenol (Gibco BRL, Ultra Pure Buffer-Saturated Phenol, pH 7.49 - 7.79) and 300 µl of chloroform/isoamyl alcohol 24:1 (Sigma). Rotate tubes for 10 minutes.

3. Centrifuge at 11,000 rpm on a microcentrifuge for 15 minutes.

4. Transfer upper, aqueous layer to a new Eppendorf tube using wide ended pipette tips.

5. Add 600 µl of chloroform/isoamyl alcohol 24:1 to each sample and rotate for 10 minutes.

6. Centrifuge at 11,000 rpm on a microcentrifuge for 10 minutes.

7. Transfer upper, aqueous layer to a new 1.5 ml Eppendorf tube using wide ended pipette tips. Add ice cold absolute ethanol - 2x volumes of transferred layer, mix, and then add 3 M sodium acetate (NaAc) - 1/10 of volume of transferred layer. Alternatively add 1 ml of ice-cold isopropanol.

8. Place in freezer at -20°C overnight (or at least 40 minutes).

9. Centrifuge for 25 minutes at 11,000 rpm.

10. Gently pour off ethanol leaving DNA pellet behind.

11. Add 400 µl of ice-cold 70% ethanol. Gently invert twice and then centrifuge at 11,000

rpm for 5 minutes. Gently pour off ethanol, or remove with a pipette.

**12.** Repeat **11.**

**13.** Air dry pellet and add 50 µl of 1xTE buffer (or elution buffer from the Qiagen DNeasy Tissue Kit (Qiagen) to resuspend the DNA pellet.

<u>Applied alterations to the above-described protocol were</u>:

- At step **1**, digestion with proteinase K was extended from 2-3 to 5 hours at 55°C, and then samples were left for further digestion at room temperature overnight. Next day, an additional 10 µl Proteinase K Solution of 15.6 mg/ml (Boehringer Mannheim) was added and digestion was carried out for further 8 hours at 55°C.

- At step **2** and **5**, the tubes were rotated for 15 minutes instead of 10 minutes.

- At step **6**, centrifugation was extended from 10 to 15 minutes.

- The modifications applied at steps **7** and **8** were by adding only 2x volume of absolute ethanol and placing the samples in the freezer (-20°C) for 48 hours. Then, 1/10 volume of 3 M NaAc (Sigma, pH 5.5) was added and samples were placed in the freezer (-20°C) for only 10 minutes.

- At step **13:** An additional placing of DNA samples (DNA re-suspended in 50 µl of 1xTE buffer) in a water bath at 55°C and 70°C in order to aid dissolving the DNA pellets.

## Protocol ($G_2$)

**The Qiagen DNeasy Tissue Kit** (Qiagen, Crawley, West Sussex, UK). A very modified version of the manufacturer's protocol - Southampton's version of the protocol was tested on formalin-fixed *Nezumia* samples. Applied protocol is as follows:

**1.** Small macerated pieces of tissue were placed into four 1.5-ml microcentrifuge tubes. Into each tube different extraction buffers were added (in two tubes were added 250 µl of the buffer ATL from the Kit, in one was added 100 µl Buffer ATL + 100 µl of extraction

buffer from the previously described protocol ($G_1$) (40 µl of 1M Tris-HCl pH 8.0 + 20 µl of 25% SDS + 350 µl $H_2O$), and in one was added only 250 µl of extraction buffer from the previously described protocol ($G_1$).

2. Then, into each tube, was added 50 µl of Proteinase K solution from the Kit (concentration 20 mg/ml) and mixed by vortexing, after which the tubes were placed in a water bath for 2 hours at 55°C - occasionally vortexing during incubation. The samples were removed from the water bath and placed at room temperature overnight for continuation of proteinase K digestion.

3. Next day, an additional 10 µl of Proteinase K (20 mg/ml) was added into each tube and placed into a water bath at 55°C for 10 hours with occasional vortexing of tubes.

4. 250 µl of AL Buffer from the Kit and 250 µl of ethanol were added, and then the samples were vigorously mixed by vortexing.

5. The mixtures from step 4 were transferred into a DNeasy mini column in a 2-ml collection tube. It was centrifuged at 8,000 rpm for 1 min. Flow-through and collection tubes were discarded.

6. The DNeasy mini column was transferred to a new 2-ml collection tube, and 500 µl of the Buffer AW1 (from the Kit) was added. After centrifugation at 8,000 rpm for 1 min, flow-through and collection tubes were discarded.

7. The DNeasy mini column was placed in a new 2-ml collection tube, and 500 µl of the Buffer AW2 (from the Kit) was added. After centrifugation at 12,000 rpm for 3 min., flow-through and collection tubes were discarded.

8. The DNeasy mini column was removed and placed in a clean 1.5-ml microcentrifuge tube. First, 100 µl (the first elution) and then 130 µl (the second elution) of AE Buffer (Elution Buffer) from the Kit were added directly on to DNeasy membrane. Samples were incubated at room temperature for 1 min, and then centrifuged at 8,000 rpm to elute.

## Protocol ($G_3$)

**The microLYSIS Kit.** A commercially available microLYSIS reagent (Microzone Ltd, West Sussex, UK), applying the manufacturer's original protocol, was also tested in the Southampton Laboratory for *Nezumia* formalin-DNA extractions. The microLYSIS method is a simple, "one step" protocol: a small macerated piece of tissue (previously only briefly washed in sterile water) was placed into a PCR tube to which was added 20 µl of microLYSIS reagent. Then, the tube was placed in a PCR thermal cycler applying one PCR cycle with the following cycling profile:

Step 1: 65°C for 5 min,

Step 2: 96°C for 2 min,

Step 3: 65°C for 4 min,

Step 4: 96°C for 1 min,

Step 5: 65°C for 1 min,

Step 6: 96°C for 30 s,

Step 7: 20°C hold.

***A phenol based DNA extraction protocol used in the Molecular Laboratory of the Natural History Museum, London (applied only on Nezumia ethanol preserved tissue used for the sequence validation) – protocol (H)***

## Protocol (H)

**Phenol-based protocol (PCI).** In the DNA Laboratory of the Natural History Museum (NHM), London, a modified phenol-based protocol (Dr David Johnston, personal communication) was applied for DNA extractions from ethanol preserved tissue samples of *Nezumia*. The protocol is outlined below:

1. Prior to DNA extraction, small pieces of tissue (15-20 mg) were washed in 1x TE buffer (pH 8.0) for 3 hours and in sd$H_2O$ for a few minutes.

2. Macerated tissue was placed into 2 ml sterile tubes. Then, 500 µl of extraction buffer (50 mM Tris-HCl (pH 8.0), 50 mM EDTA, 100 mM NaCl, 1% SDS) and 10 µl of Proteinase K (20 mg/ml in water solution) were added. The proteinase K digestion was carried out in the Hybaid mini hybridisation oven at 55°C, with tubes rotating overnight.

3. Next day, an additional 10 µl of Proteinase K (20 mg/ml; Sigma) + 10 µl of RNase A (10 mg/ml; Sigma) were added into the tubes, and samples were then again placed at 55°C for few hours in order to complete digestion.

4. When proteinase K digestion was completed, phenol (500 µl; Sigma) was added into each tube. The tubes were gently inverted (at least 20 times) to mix the phenol with the digested tissue sample.

5. Then, samples were centrifuged at 13,000 rpm for 5 minutes at room temperature. The supernatant (SN) with DNA was carefully removed and transferred into a new tube using wide-open pipette tips.

6. A mixture of 250 µl of phenol + 250 µl chloroform/isoamyl alcohol (24:1) was added to the SN, and the tubes were gently mixed by inversion. Centrifugation, 13,000 rpm for 5 minutes at room temperature, was applied.

7. The supernatant (SN) with DNA was transferred into new tubes and then, 500 µl of chloroform/isoamyl alcohol (24:1) was added. After gentle mixing, samples were centrifuged at 13,000 rpm for 5 minutes. The SN was transferred into new tubes using wide-open tips.

8. DNA was precipitated by adding 2 volumes of absolute ethanol and 1/10 volume of 3 M NaAc (pH 5.2). DNA precipitation was carried out in a fridge at 4°C (for 45 minutes, or overnight). Then, the samples were centrifuged at 13,000 rpm for 5 minutes at room temperature. The supernatant was carefully decanted.

9. The pellet was washed only once by adding 1 ml of 70% ethanol and applying short vortexing. Then, the samples were centrifuged at 13,000 rpm for 5 minutes at room temperature. The supernatant (ethanol with impurities) was carefully decanted. In order to remove any residues of ethanol, drops on the tube's sides were aspirated using a Pasteur pipette, or vacuum pump.

10. The pellet was dried for a few minutes at 55°C, or for 20-30 min at 37°C and then dissolved in 100 µl of ultrapure (HPLC) sd$H_2O$. In order to dissolve fully the DNA pellet, before storage at -20°C, the samples were kept at room temperature for 2 hours, or at 37°C for 30 min, or overnight at 4°C.

### 2.2.2.1. *DNA extraction protocols tested on supplementary specimens – ten differently preserved specimens of mackerel*

Pre-extraction treatments:

**A -** Tissue was washed overnight in 1xTE buffer at room temperature prior to DNA extraction.

**B -** Tissue was washed for a few minutes in $sdH_2O$, and then dried for a few minutes at room temperature prior to DNA extraction.

**C -** Tissue was washed twice for 5 minutes in PBS (phosphate-buffered saline) and twice in $sdH_2O$ for a couple of minutes, and then briefly dried for a few minutes at room temperature prior to DNA extraction.

**D -** Tissue was washed in 1xGTE buffer for 24 hours (gently rotating the samples) at room temperature and frequently changing the buffer, then washed in $sdH_2O$ for 10-15 minutes, and then briefly dried for a few minutes at room temperature prior to DNA extraction.

Protocols:

- **Protocol (A)** developed in the study with *Nezumia*, but with a modification is that liquid nitrogen was used instead of dry ice. It was applied pre-extraction treatment D.

- **Promega Wizard Genomic DNA Purification Kit** (Promega, Madison, USA) was applied following the manufacturers` instructions for isolation of genomic DNA from animal tissue (mouse tail) – **protocol (F).** It was applied pre-extraction treatment B.

- **Phenol-based protocol (H)** as described in the previous section (precipitation with sodium acetate), or with the modification that DNA precipitation was performed with polyacrylamide instead with sodium acetate. It was applied pre-extraction treatment A.

- **"Amersham Biosciences" GenomicPrep Cells and Tissue DNA Isolation Kit** (Protocol: Extraction of DNA from animal tissue; Amersham Biosciences UK Limited) was applied following the manufacturers` instructions. It was applied pre-extraction treatment B.

- **Qiagen DNeasy Tissue Kit,** (Protocol: Isolation of Genomic DNA from formalin-fixed tissue; Qiagen, Crawley, West Sussex, UK) was applied following the manufacturers` instructions. It was applied pre-extraction treatment C.

### *2.2.2.2. Quantification of extracted DNA*

The Saran Wrap Method (Sambrook *et al.* 1989) was applied for ethidium bromide quantification of DNA samples extracted from different fish species, specimens and tissues. The adapted protocol is outlined below:

1. Stretch a sheet of Saran Wrap over an ultraviolet transilluminator.
2. Spot series of 3 µl volume of 1xTE buffer containing 2 µg/ml ethidium bromide.
3. Add to each spot equal volumes (3 µl) of a series of DNA concentration standards (0, 1, 2.5, 5, 10 and 20 µg/ml) in an ordered array on the Saran Wrap and mix by pipetting up and down. The standard DNA solutions for this project contained DNA of salmon testes (Sigma).
4. Then, add to each spot (containing TE buffer with 2 µg/ml ethidium bromide) 3 µl of the DNA sample on to the Saran Wrap.
5. Photograph the spots using short-wavelength ultraviolet illumination. Estimate the concentration of DNA by comparing the intensity of fluorescence in the sample with those of the standard DNA solutions.

During the study, there were attempts to use spectrophotometric and gel molecular weight marker measurement of DNA concentration, but degraded DNA at low concentration made these attempts unsuccessful.

A quantitative size gel marker measurement of DNA concentrations was applied for control, ethanol preserved specimens of *Nezumia* – HyperLadder I and HyperLadder IV (Bioline, London, UK).

### 2.2.3. PCR amplifications and molecular investigations

Different kinds of molecular techniques and methods (PCR, RFLP, cloning, sequencing, Southern blotting, etc.) were used in this study.

#### *2.2.3.1. PCR protocols and primers*

The study mainly used three PCR marker systems:

**(1)** Random amplified polymorphic DNAs (RAPDs),

**(2)** Universal mitochondrial PCR primers for targeted mitochondrial genes (mtDNA),

**(3)** RAPD-derived, *Nezumia*-specific PCR primers developed and designed in this project.

A technical preparation of PCR reactions and their handling were the same for all PCR markers. All reagents and PCR mastermix were kept on ice during PCR set ups. PCR reactions were loaded directly from ice into 94°C preheated PCR cycler block. PCR reactions were not loaded into the end of rows of a PCR thermocycler block due to possible unequal heating on the edges of the block during cycling. Negative controls of PCR reactions were carried out with each PCR experiment.

Thermal cyclers used/tested in this research:

- Hybaid Omn-E thermal cycler (0.2 ml tubes) (Hybaid Ltd., Middlesex, UK) – ***used as the main thermocycler during the study (Nescot Laboratory)***;
- Hybaid Omn-E thermal cycler (0.5 ml tubes) – tested in the Nescot Laboratory;
- Hybaid PCR Express thermal cycler with a cooling system (0.2 ml tubes) – tested in the Nescot and Southampton's laboratories;
- Hybaid Omnigene thermal cycler without hot lid – tested in the Nescot Laboratory;
- MWG-Biotech Primus thermal cycler with a cooling system (0.2 ml tubes) – tested in the Nescot Laboratory;
- Perkin Elmer thermal cyclers with a cooling system (0.2 ml tubes) – Types: DNA Thermal Cycler 480 (Southampton`s Laboratory) and GeneAmp PCR System 2400 (NHM Laboratory);
- GeneAmp PCR System 2700 (Applied Biosystems) – used in the NHM Laboratory.

*Taq* polymerases used/tested in this research:

- AGS*Gold* DNA Polymerase (Hybaid Ltd., Middlesex, UK) – **used as the main *Taq* polymerase during the study (Nescot Laboratory);**
- Promega *Taq* polymerase – tested in the Nescot Laboratory;
- Red Hot DNA Polymerase (Advanced Biotechnologies Ltd., Surrey, UK) – tested in the Nescot Laboratory;
- Pharmacia *Taq* polymerase (Pharmacia Biotech, USA) – tested in the Nescot Laboratory;
- Perkin Elmer Ampli*Taq* polymerase – tested in the Southampton's Laboratory;
- Qiagen *Taq* DNA Polymerase Kit – tested in the Southampton's Laboratory;
- pureRe*Taq* Ready-To-Go PCR Beads (Amersham Biosciences UK Ltd., Buckinghamshire, UK) – tested in the NHM Laboratory;
- AB Thermoprime Plus DNA Polymerase; ABgene 2x PCR Master Mix (1.5mM $MgCl_2$) – ready-to-use master mix (ABgene, Surrey, UK) – NHM Laboratory.

Hybaid Omn-E thermocycler (for 0.2 ml tubes) and Haybaid-AGS*Gold* DNA polymerase were used throughout the project as the main thermal cycler and *Taq* polymerase. Others were only tested in a few experiments in order to compare for differences in results, or if specific equipment/material was available in a particular laboratory.

#### 2.2.3.1.1. RAPD-PCR experiments

RAPD-PCR amplifications were regularly carried out at three different PCR-DNA template concentrations of the same DNA extract. PCRs were performed in duplicate for each PCR-DNA template concentration. Relative PCR-DNA template concentrations:

- 1x PCR-DNA template concentration (0.5 µl of DNA suspension in 25 µl volume of PCR reaction)
- 5x PCR-DNA template concentration (2.5 µl of DNA suspension in 25 µl volume of PCR reaction)
- 10x PCR-DNA template concentration (5 µl of DNA suspension in 25 µl volume of PCR reaction).

**For RAPD-PCR** amplifications, three kit sets were used: AB-0320-Kit2, AB-0320-Kit7 and AB-0320-Kit8 (Advanced Biotechnologies, Surrey, UK). Each kit contains twenty 10-base oligonucleotide primers with randomly selected bases (see sequences of RAPD primers in Table 1).

**Table 1 -** Sequences of RAPD primers used in this study (Advanced Biotechnologies, Surrey, UK)

| # | Primer name | Sequence | # | Primer name | Sequence |
|---|---|---|---|---|---|
| | ***Kit 2*** | | | | |
| 1. | AB2-01 | 5`-CCCAAGGTCC-3` | 11. | AB2-11 | 5`-GAGTCTCAGG-3` |
| 2. | AB2-02 | 5`-GGTGCGGGAA-3` | 12. | AB2-12 | 5`-TTATCGCCCC-3` |
| 3. | AB2-03 | 5`-CCAGATGCAC-3` | 13. | AB2-13 | 5`-CCCGATTCGG-3` |
| 4. | AB2-04 | 5`-GTGACATGCC-3` | 14. | AB2-14 | 5`-TGCGGCTGAG-3` |
| 5. | AB2-05 | 5`-TCAGGGAGGT-3` | 15. | AB2-15 | 5`-ACGCACAACC-3` |
| 6. | AB2-06 | 5`-AAGACCCCTC-3` | 16. | AB2-16 | 5`-GGTGACTGTG-3` |
| 7. | AB2-07 | 5`-AGATGCAGCC-3` | 17. | AB2-17 | 5`-CTACTGCCGT-3` |
| 8. | AB2-08 | 5`-TCACCACGGT-3` | 18. | AB2-18 | 5`-GGACTGCAGA-3` |
| 9. | AB2-09 | 5`-CTTCACCCGA-3` | 19. | AB2-19 | 5`-ACGGCGTATG-3` |
| 10. | AB2-10 | 5`-CACCAGGTGA-3` | 20. | AB2-20 | 5`-AACGGTGACC-3` |
| | ***Kit 7*** | | | | |
| 21. | AB7-01 | 5`-CAAAGGGCGG-3` | 31. | AB7-11 | 5`-CAATCGGGTC-3` |
| 22. | AB7-02 | 5`-CTGAACCGCT-3` | 32. | AB7-12 | 5`-AAGAGGGCGT-3` |
| 23. | AB7-03 | 5`-TCTCGCCTAC-3` | 33. | AB7-13 | 5`-GGTTCCTCTG-3` |
| 24. | AB7-04 | 5`-GTAGGCCTCA-3` | 34. | AB7-14 | 5`-GAACGAGGGT-3` |
| 25. | AB7-05 | 5`-ACCGCATGGG-3` | 35. | AB7-15 | 5`-TTTGCCCCGT-3` |
| 26. | AB7-06 | 5`-AAGTGCACGG-3` | 36. | AB7-16 | 5`-AACGGGCGTC-3` |
| 27. | AB7-07 | 5`-CCCTACTGGT-3` | 37. | AB7-17 | 5`-GGCAAACCCT-3` |
| 28. | AB7-08 | 5`-GGCAGGCAAG-3` | 38. | AB7-18 | 5`-ACGAGAGGCA-3` |
| 29. | AB7-09 | 5`-TCGCTTCTCC-3` | 39. | AB7-19 | 5`-CTTGGCACGA-3` |
| 30. | AB7-10 | 5`-AAGAGGCCAG-3` | 40. | AB7-20 | 5`-TCTTCGGAGG-3` |
| | ***Kit 8*** | | | | |
| 41. | AB8-01 | 5`-GGCATCGGCT-3` | 51. | AB8-11 | 5`-ACGGCGATGA-3` |
| 42. | AB8-02 | 5`-AGCCGTTCAG-3` | 52. | AB8-12 | 5`-GACGCGAACC-3` |
| 43. | AB8-03 | 5`-GGGTCCAAAG-3` | 53. | AB8-13 | 5`-ACGCTGCGAC-3` |
| 44. | AB8-04 | 5`-CTATCCTGCC-3` | 54. | AB8-14 | 5`-TGGTGCACTC-3` |
| 45. | AB8-05 | 5`-GTCGTAGCGG-3` | 55. | AB8-15 | 5`-GACACAGCCC-3` |
| 46. | AB8-06 | 5`-TGCCGCACTT-3` | 56. | AB8-16 | 5`-AAGGCACGAG-3` |
| 47. | AB8-07 | 5`-ACGAGCATGG-3` | 57. | AB8-17 | 5`-CCTCACGTCC-3` |
| 48. | AB8-08 | 5`-AAGCCCCCCA-3` | 58. | AB8-18 | 5`-TCGCGGAACC-3` |
| 49. | AB8-09 | 5`-TCGCTGGTGT-3` | 59. | AB8-19 | 5`-GGCAAAGCTG-3` |
| 50. | AB8-10 | 5`-TCGGGGCATC-3` | 60. | AB8-20 | 5`-CCTGTTCCCT-3` |

For RAPD amplifications, the Hybaid Omn-E thermal cycler (0.2 ml tubes) was programmed as follows: initial denaturation at 94°C for 1 min, then 45 cycles of: 5 s at 94°C, 30 s at 36°C, and 1 min at 72°C. One cycle of a final extension step of 2 min at 72°C was carried out. A RAPD-PCR reaction mix contained: 0.5 µl, 2.5 µl, or 5 µl of template DNA suspension; 1 x DNA Polymerase reaction buffer (Buffer GOLD 2 for AGSGold DNA Polymerase containing 75 mM Tris-HCl (pH 90), 20 mM $(NH_4)_2SO_4$, 0.01% Tween-20); 2.5 mM $MgCl_2$; 0.4 mM dNTPs mix; 200 nM primer; 1 unit AGS*Gold*™ *Taq* Polymerase (Hybaid, Middlesex, UK) in a final reaction volume of 25 µl. This was used as the main RAPD-PCR set up throughout the project.

### 2.2.3.1.2. Mitochondrial PCR protocol

Throughout the project, six sets of mitochondrial primers were used for amplifications of targeted mitochondrial sequences (Table 2).

**Table 2 -** List of mitochondrial primers used in the *Nezumia* study

| *Set* | **Primer name** | *Sequence of primers* | *PCR product size* |
|---|---|---|---|
| *1)* | **Cytochrome b (cyt b):** | | |
| | [1] *L 14841* | **5`-CCAACATCTCAGCATGATGAAA-3`** | |
| | [2] *H 15163* | **5`-TGAGGACAAATATCATTCTGAG-3`** | **351 bp** |
| *([1] universal - from: DeSalle et al. 1993b; [2] Hammond 1997, personal communication)* | | | |
| *2)* | **Cytochrome oxidase subunit II (COII):** | | |
| | [1] *L 7450* | **5`-AAAGGAAGGAATCGAACCCCC-3`** | |
| | [2] *H 8055* | **5`-AAGACGTCCTCCACTCATGAGC-3`** | **605 bp** |
| *([1] for fish - from: DeSalle et al. 1993b; [2] Hammond 1997, personal communication)* | | | |
| *3)* | **Cytochrome oxidase subunit III (COIII):** | | |
| | [1] *L 9459* | **5`-TTATTTATTGCATCAGAAGT-3`** | |
| | [2] *H 9924* | **5`-TCAACAAAGTGTCAGTATCA-3`** | **465 - 500 bp** |
| *([1.2] universal - from: DeSalle et al. 1993b)* | | | |
| *4)* | **Set for 12S rRNA:** | | |
| | [1] *MIT12SAI5* | **5`-AAACTAGGATTAGATACCTATTAT-3`** | |
| | [2] *MIT12SB3* | **5`-AAGAGCGACGGGCGATGTGT-3`** | **433 bp** |
| *([1.2] Hammond 1997, personal communication)* | | | |

*5)* **Set for COI + tRNA +COII:**

| | | |
|---|---|---|
| [1]*MITLEP2F* | 5`-ATTATGTTCTTTCTATAGG-3` | |
| [2]*MITLEP2R* | 5`-AAAATGTTAAATTTACTCC-3` | **146 bp** |

*([1,2] Hammond 1997, personal communication)*

*16S)* **Set for 16S rRNA:**

| | | |
|---|---|---|
| [1]*16S AR* | 5`-CGCCTGTTTATCAAAAACAT-3` | |
| [2]*16S BR* | 5`-CCGGTTTGAACTCAGATCATG-3` | **570 bp** |

*(universal: [1]Kocher et al. 1989; [2]Palumbi et al. 1991)*

---

The thermal cycling was as follows: initial denaturation at 94°C for 1 min, then 35 - 40 cycles of 30 - 45 s at 94°C, annealing temperature at 38 - 51°C for 30 - 60 s depending on the type of mitochondrial primer set used, and extension 1 - 2.45 min at 72°C. A final extension step of 2 - 7 min at 72°C was carried out. The best performance for mitochondrial amplifications and the program mostly used on the Hybaid Omn-E thermal cycler was:

STEP 1: 94°C for 1 min (initial denaturation)
then 40 cycles of
STEP 2: 94°C for 45 s (denaturation)
STEP 3: 40°C for 1 min (annealing)
STEP 4: 72°C for 2 min 45 s (extension)
and then 1 cycle of
STEP 5: 72°C for 7 min (final extension)

Amplification reactions were performed in volumes of 25 µl containing: 2.5 µl of 10 x reaction buffer (Hybaid-AGS GOLD 2 buffer for Hybaid-AGSGold DNA Polymerase containing 750 mM Tris-HCl (pH 90), 200 mM $(NH_4)_2SO_4$, 0.1% Tween-20), 2.5 µl of 25 mM $MgCl_2$ (from the Hybaid PCR Kit), 100 µM of each dNTP (Pharmacia Biotech, or Hybaid), 1.25 µl of each primer (5 pmol/µl), 1.25 units *Taq* polymerase (Hybaid-AGS*Gold* DNA Polymerase), and 2.5 - 10 µl of template DNA (absolute concentration of DNA in PCR reaction varied – depending on a particular DNA extract and if a diluted DNA suspension was used).

The PCR amplifications of the mitochondrial 16S gene on DNA extracted from formalin-fixed specimens of *Nezumia* and DNA of ethanol preserved *Coryphaenoides* were carried out by using Hybaid PCR Express and Perkin Elmer DNA Thermal Cycler 480 in the

Molecular Ecology Laboratory of the Southampton Oceanography Centre. PCR thermal cycling consisted of 40 cycles of denaturation for 45 s at 94°C, annealing 1 min at 45°C, and extension for 3 min at 72°C. The final extension was 7 min at 72°C. PCRs were carried out with 1-6 µl DNA extract, 2 µl of 10x reaction buffer (Perkin Elmer [PE] PCR Buffer II, or Qiagen *Taq* polymerase buffer), 2.4 µl 25mM $MgCl_2$, 1.6 µl of dNTPs mix (PE; 10 mM), 1 µl of each primer (10 pmol/µl), and 1 µl of *Taq* polymerase (1 U/µl; Perkin Elmer Ampli*Taq* DNA Polymerase) in a volume of 20 µl. Qiagen *Taq* polymerase was tested in a few PCR experiments. Qiagen Q-solution (3 µl of 5x Q-solution) from Qiagen PCR kit (contains glycerol) was applied in most PCRs in order to improve a performance of the mitochondrial PCR amplifications.

PCR amplifications of the mitochondrial 16S gene were carried out on DNA of ethanol preserved tissue samples of *Nezumia* and ten differently preserved specimens of mackerel using a PE GeneAmp PCR System 2400, or GenAmp PCR System 2700, thermal cycler at the Molecular Laboratory of the Natural History Museum, London. PCR reactions were prepared in a volume of 10 µl and 25 µl using ABgene ready *Taq* PCR mix (2x PCR master mix with Thermoprime Plus DNA polymerase; according to manufacturer's information, containing in a final reaction volume: 1.25 units of *Taq*, 75 mM Tris-HCl, 20 mM $(NH_4)_2SO_4$, 1.5 mM $MgCl_2$, 0.01% Tween 20, 0.2 mM of each dNTP). If puRe*Taq* Ready-To-Go PCR Beads (Amersham Biosciences UK Limited, Buckinghamshire, UK) were used, PCRs were performed in 25 µl of a PCR reaction volume containing ~2.5 units of puRe*Taq* DNA polymerase, 10 mM Tris-HCl (pH 9 at room temperature), 0.2 mM of each dNTP (dATP, dCTP, dGTP and dTTP), 50 mM KCl, 1.5 mM $MgCl_2$, and stabilizers, including BSA (according to the manufacturer's product sheet). PCR reactions were carried out with 2.5 µl of each primer (10 pmol/µl) and 5 µl DNA extract (or 10 ng of *Nezumia* template DNA) in a volume of 25 µl, i.e. with 1 µl of each primer (10 pmol/µl) and 2 µl DNA extract in a volume of 10 µl. PCR thermal cycling consisted of an initial denaturation at 94°C for 1 min, followed by 40 cycles of denaturation for 30 s at 94°C, annealing for 30 s at 48°C, and extension for 1 min at 72°C. A final extension step of 7 min at 72°C was carried out.

Another set of mitochondrial primers for cytochrome b (developed by Finnerty and Block (1992) for blue marlin, *Makaira nigricans*) was tested in the NHM Laboratory on ethanol preserved samples of *Nezumia* and ten differently preserved mackerel specimens:

L15079 5`-GAGGCCTCTACTATGGCTCTTACC-3`

H15497 5`-GCTAGGGTATAATTGTCTGGGTCGCC-3`.

Expected size of PCR product for *Scomber* is 418 bp (Scoles *et al.* 1998).

Primers for cytochrome b (Set 1), cytochrome oxidase subunits II (Set 2), cytochrome oxidase subunits III (Set 3) and 16S were synthesized by MWG-Biotech (MWG-Biotech UK Ltd., Milton Keynes, UK), whereas primers for Set 4 (12S) and Set 5 were synthesized by Pharmacia Biotech. Mitochondrial primers used in the NHM Laboratory on ethanol preserved specimens of *Nezumia* and ten differently preserved specimens of mackerel (cyt b for blue marlin and 16S) were synthesized by Sigma-Genosys (Haverhill, UK). All primers were delivered lyophilised. The stock solution of primers (concentration of 100 pmol/µl) was prepared with $sdH_2O$.

### 2.2.3.1.3. PCR protocol for newly designed *Nezumia*-specific, RAPD-derived, primers

Designed oligonucleotide primers were synthesized by MWG-Biotech (MWG-Biotech UK Ltd., Milton Keynes, UK) and by Sigma-Genosys (Sigma-Genosys Ltd., Haverhill, UK). Lyophilised primers were dissolved in sterile water (at a stock concentration solution of 100 pmol/µl) and kept at –20°C. A working stock solution was prepared at a concentration of 5 and 10 pmol/µl.

The Hybaid Omn-E thermal cycler with hot lid was used for these PCR amplifications in the Nescot Laboratory. PCR cycle conditions were: 1 cycle of 94°C for 1 min (the initial denaturing); 40 cycles of 94°C for 45 s (denaturing), 40°C for 1 min (annealing), 72°C for 2 min 45 s (extension), with a final extension phase of 72°C for 7 min (1 cycle). Different annealing conditions, i.e. duration 30 s to 2.5 min and temperature 40-55°C were tested, but the best performance was with an annealing step of 40°C for 1 min. PCRs were carried out with 0.5 µl, 2.5 µl, or 5.0 µl suspension of extracted total cellular DNA, 2.5 µl of 10 x reaction buffer (Hybaid-AGS GOLD 2 buffer), 2.5 µl of 25 mM $MgCl_2$, 400 µM of dNTP mix (Pharmacia Biotech, or Hybaid), 1.25 µl of each primer (5 pmol/µl), 1.25 units of Hybaid-AGS*Gold Taq* DNA polymerase in a volume of 25 µl.

For DNA extracted from ethanol-preserved tissue samples of *Nezumia aequalis* and *Nezumia* cf. *micronychodon*, PCRs were performed in the Molecular Laboratory of the

Natural History Museum, London by using a GeneAmp PCR System 2700 (Applied Biosystems) and a PE GeneAmp PCR System 2400 thermal cyclers, and puRe*Taq* Ready-To-Go PCR Beads (Amersham Biosciences) and ABgene ready PCR mix (ABgene, Epsom, Surrey, UK) PCR kits. A content of PCR reactions was the same as for mitochondrial PCRs performed in the NHM (see p 94). PCR reactions were subjected to the following thermal cycles: 1 cycle at 94°C for 1 min of the initial denaturing; 40 cycles at 94°C for 15 - 45 s denaturing, 40°C - 48°C for 15 s - 1 min 45 s annealing, 72°C for 30 s - 2 min 45 s extension; and 1 cycle at 72°C for 7 min final extension. Different durations of annealing (15 s – 1 min 45 s) and annealing temperatures (40 - 48°C) were tested, but the best performance was 48°C for 15 s and 30 s for the annealing step. Also, different lengths (30 s – 2 min 45 s) of extension temperature (72°C) were tested, but the best performance was 72°C for 30 s of annealing step. To summarise, the best performance for ethanol-DNA of *Nezumia* and the program mostly applied was:

STEP 1: 94°C for 1 min (initial denaturising)
then 40 cycles of
STEP 2: 94°C for 15 s (denaturation)
STEP 3: 48°C for 15 s (annealing)
STEP 4: 72°C for 30 s (extension)
and then 1 cycle of
STEP 5: 72°C for 7 min (final extension)
STEP 6: 4°C for ~ (hold)

Selected primer sets were also tested against DNA extracted from fresh/frozen tissue of mackerel (*Scomber scombrus*, Family Scombridae) applying the same PCR conditions as for ethanol-DNA of *Nezumia* (tested in the NHM).

#### 2.2.3.1.4. Other primers and PCR experiments

The universal **ITS primers** were tested in a few PCR amplifications with DNA from formalin-fixed *Nezumia* specimens. The primers used in these PCR experiments are:

ITS-1: 5` - TCC GTA GGT GAA CCT GCG G - 3`, and
ITS-4: 5` - TCC TCC GCT TAT TGA TAT GC - 3`
(Allainguillaume 2000, personal communication).

The Hybaid Omn-E thermal cycler cycling conditions were: 1 cycle at 94°C for 1 min initial denaturing; 40 cycles of 94°C for 45 s (denaturing), 40°C for 1 min (annealing), 72°C for 2 min 45 s (extension); and 1 cycle of 72°C for 7 min (final extension). PCRs were carried out with 0.5 µl, 2.5 µl, or 5.0 µl extract DNA, 2.5 µl of 10 x reaction buffer (Hybaid-AGS GOLD 2 buffer), 2.5 µl of 25 mM $MgCl_2$, 400 µM of dNTP mix (Pharmacia Biotech, or Hybaid), 1.25 µl of each primer (5 pmol/µl), 1.25 units of Hybaid-AGS*Gold Taq* DNA polymerase in a volume of 25 µl.

Primers for the two **microsatellite** loci (CR 1/16 and CR 2/40), developed for *Coryphaenoides* by the Ecological Molecular Group of the Oceanography Centre, University of Southampton, UK (Alex Rogers 1999, personal communication) were tested on DNAs extracted from formalin-fixed specimens of *Nezumia* and ethanol preserved samples of *Coryphaenoides*.

The primers for the **28S gene D2 domain** and for the rhodopsin gene were tested only on DNA extracted from the ethanol preserved muscle tissue of *Nezumia*, and on DNA extracted from ten differently preserved tissue samples of mackerel (*Scomber scombrus*). Primers used to amplify 28S gene D2 were developed by Lecointre *et al.* (1997) for fish:

C`1: **5` - ACC CGC TGA ATT TAA GCA T - 3`**, and

D2: **5` - TCC GTG TTT CAA GAC GGG - 3`**

Primers used to amplify the **rhodopsin** gene were developed by Zaragueta-Bagils *et al.* (2002) for fish:

Rh545 (forward): **5` - GCA AGC CCA TCA GCA ACT TCC G - 3`**, and

Rh667 (reverse): **5` - AYG AGC ACT GCA TGC CCT - 3`**

The **28S and rhodopsin genes** were amplified using the same PCR conditions as for 16S mitochondrial PCRs performed in the NHM (see p 94).

### 2.2.3.2. *Identification of PCR products by electrophoresis*

PCR products were visualized on 1% agarose gel using ethidium bromide staining and UV transillumination. Five microlitres of loading buffer (25 % Ficoll and 0.025 % bromophenol blue) was added to 15 µl PCR reaction product and electrophoresis was run at 6 V/cm for 1.5 hours. In Southampton and in the NHM Laboratory, a slightly different gel loading buffer was used (0.25 % bromophenol blue, 0.25 % xylene cyanol, 30 % glycerol, and 69.5 ml $dH_2O$ – in Southampton, that is, 20% sucrose, XC/BP Blue – 6x Loading dye in the NHM) and a smaller amount of amplification reaction for examination (5-10 µl). Tris-borate buffer (Tris-borate/EDTA buffer) was used as electrophoresis buffer in the Nescot Laboratory, whereas 1x TAE buffer (Tris base/ glacial acetic acid/ EDTA buffer) was used in the Southampton and the NHM laboratories.

As DNA size markers, 1 Kb DNA Ladder (Gibco BRL, or Promega) was mainly used. A quantitative HyperLadder I and HyperLadder IV (Bioline) was used in the NHM Laboratory.

Photographs were taken by Polaroid camera (Polaroid CU-5 Land Camera) with a red filter using black and white Polaroid 667 film. In the Southampton and NHM laboratories, the gel was photographed and recorded using the gel-documentation system (Uvidoc in Southampton; UVP at the NHM).

#### 2.2.3.2.1. Extraction of DNA fragments from gel slices

DNA of selected RAPD-PCR fragments were purified from gels in order to be re-amplified, cloned, or labelled and used for Southern blotting. After running PCR amplification products on the agarose gel, fragments were excised from the gel by sterile scalpel. DNA extractions from excised gel slices were carried out immediately, or gel slices were stored at -20°C until needed. The Hybaid Recovery Gel Extraction Kit (Hybaid Ltd., Middlesex, UK) was used following the manufacturer's instructions.

### *2.2.3.3. Southern blotting*

Selected DNA RAPD-PCR fragments were used for DIG labeling and Southern blotting.

#### 2.2.3.3.1. DIG labelling

DNA RAPD-PCR fragments (cloned and/or PCR amplicons) were labelled by digoxigenin-dUTP (DIG) using DIG DNA Labelling and Detection Kit (Boehringer Mannheim) according to the manufacturer's protocol.

#### 2.2.3.3.2. Southern blotting/ hybridisation

The method of overnight capillary transfer of the nucleic acid to the membrane was used as described in Hammond (1995, 1997):

Transfer of gel DNA (RAPD-PCR DNA fragments) on to the membrane:

1) The gel with electrophoretically separated RAPD-PCR fragments (bands) was immersed in NaOH/ Saline buffer (0.5 M of NaOH, 1.5 M NaCl) for 20 minutes. The dish with the gel was gently rocked at intervals.

2) Then, the gel was transferred into Tris/ Saline buffer (0.5 M Tris-HCl; 3 M NaCl, pH 7.4) and incubated for 15 minutes with intermittent agitation.

3) Two pieces of Perspex were wrapped around with a piece of Whatman 3MM filter paper and placed in a dish. The 10 x SSC buffer (Saline sodium citrate containing: 1.5 M NaCl and 0.15 M sodium citrate pH 7.0) was added almost to the top of the Perspex. This "construction" was smoothed with a clean glass rod to remove air bubbles.

4) The gel was turned over so that the wells were face down, and it was placed centrally on the filter paper wick. A nylon membrane (Duralon-UV; Stratagene) was cut into a piece of about 2 mm smaller than the gel. The membrane was placed onto the gel so that about 1 mm of gel was showing round the edges. The cling film was placed around the gel/ membrane combination so that only membrane was exposed. Two sheets of 3 MM filter paper of the same size as the membrane were cut and wetted with 2 x SSC. It was placed on top of the membrane ensuring that there were no air bubbles. Then, two sheets of dry

3 MM paper were placed on the top. A stack of paper towels was placed on the top of the filter paper. A glass plate was placed on the top of the towels and it was compressed with a weight.

5) This "construction" allowed an overnight flow of 10 x SSC from the reservoir through the gel and membrane to the paper towels, allowing the DNA fragments to be carried from the gel onto the membrane.

6) Next day, after the transfer was completed, paper towels and 3 MM papers were discarded. The position and orientation of the gel were marked on the membrane using a pencil. The membrane was placed on the clean surface of the UV transilluminator and exposed to UV light for 5 minutes in order to bind DNA onto the nylon membrane.

A pre-hybridisation was carried out immediately, or blots (membranes with transferred DNA) were wrapped in cling film and stored in a fridge at 4°C until required for pre-hybridisation and hybridisation.

<u>Pre-hybridisation and hybridisation:</u>

1) The blot was pre-hybridised by placing into a hybridisation tube with 10 ml of the hybridisation buffer (1% w/v Blocking reagent from 10% blocking stock solution supplied with the Boehringer Mannheim hybridisation Kit; 0.1% w/v N-lauroylsarcocine; 0.02% SDS). The tube with blot was placed on the rotisserie in the hybridisation oven (Hybaid) and incubated at 68°C for 1 hour with the rotisserie turning.

2) 10 µl of prepared DIG-labelled probe was transferred into a fresh micro-centrifuge tube. Placing it in a boiling water bath for 3 minutes, the DNA probe was denatured. The probe was immediately pipetted and transferred into 2.5 ml of fresh hybridisation buffer.

3) The hybridisation tube containing the blot was taken out of the oven and the pre-hybridisation solution poured out. Then, the hybridisation solution containing the probe was added to the tube and placed again in the hybridisation oven. The incubation was carried out overnight at 68°C with rotation.

4) The membrane was removed from the tube and the hybridisation solution was discarded.

Then, 50 ml of the washing solution (2x SSC; 0.1 % SDS) was added. It was gently agitated for 5 minutes at room temperature, and then the wash solution was discarded. The wash step was repeated once more.

**5)** 50 ml of the stringent wash solution (low salt: 0.1x SSC and 0.1 % SDS, or 0.05 x SSC and 0.1 % SDS), preheated at 68°C, was added to the membrane. The tube with the membrane was placed in the oven and rotated for 15 min at 68°C. The stringent wash solution was discarded and washing step repeated.

Immunological detection of the probe on the Southern blot was carried by the antibody solution supplied with the Boehringer hybridisation kit, following manufacturer's instruction (Boehringer Mannheim).

### *2.2.3.4. Cloning experiments*

Selected RAPD-PCR fragments were excised from the gel, purified (using the Hybaid Recovery Gel Extraction Kit according to the manufacturer's protocol) and used for cloning experiments. Before proceeding to cloning, extracted gel-DNA was always run (3-5 µl) on a 1-2% agarose gel for size and DNA concentration inspection.

The pUC18 plasmid vector was used as a cloning vector for these experiments. A pUC18 vector used in these experiments was mainly obtained by minipreps from *E. coli* cultures by using Hybaid Recovery Quick Mini Spin Kit (Hybaid), or Wizard Minipreps DNA Purification System (Promega). Occasionally, ready pUC18 plasmid (0.25 µg/µl; Gibco BRL, Life Technologies Inc.) was used in these experiments. It was applied the T/A cloning method described in Hillis *et al.* (1996).

A vector preparation: digestion of the pUC18 with *Sma* I and preparation of T-tailed vector (T-overhang) - Hillis *et al.* (1996) modified protocol:

1. Digestion of the pUC18 with *Sma* I (Advanced Biotechnologies 10 U/µl, or Promega 12 U/µl) was carried out for 2 hours at 25°C. Usually, digestion was performed in 30 µl of a reaction volume. If a larger volume of a digestion reaction (50 µl – 100 µl) was applied, the digestion of vector would be extended from 2 hours to 24 hours.

- 15 µl of plasmid-vector DNA (pUC18)
- 11 µl $sdH_2O$
- 3 µl multi-core buffer (10x buffer)
- 1 µl *Sma* I .

Total reaction volume: 30 µl

2. The termination of digestion process (i.e. the inactivation of the *Sma* I restriction enzyme) was carried out by heating of the reaction sample at 55°C for 5-20 min. Digested vector was purified by precipitation with 2.5x volume of absolute ethanol and 1/10 volume of 2 M NaCl overnight at –20°C. To recover the precipitate, the sample was centrifuged for 10 min at 13,000 rpm. The pellet was dried and then re-suspended in $sdH_2O$.

3. To add T-overhang, the following were used:

45 µl DNA of digested vector (pUC18/ *Sma* I)
5.1 µl of *Taq* 10x buffer (Hybaid AGS*Gold Taq* buffer containing 15 mM $MgCl_2$)
1 µl dTTP (100 mM; pH 7.5)
0.75 µl *Taq* polymerase (5 U/ µl; Hybaid, AGS*Gold* DNA polymerase)
Total reaction volume: 51.85 µl

The incubation was carried out in a thermal cycler (Hybaid Omn-E with hot lid) at 70°C for 2 hours.

4. PCI (Phenol : Chloroform : Isoamyl Alcohol, 25:24:1) and CI (Chloroform : Isoamyl Alcohol, 24:1) extraction was carried out:
   (i) An equal volume of PCI was added, gently mixed with sample and incubated at room temperature for 5 min. Phase separation was obtained by centrifuging the sample at 13,000 rpm for 5 min. The upper aqueous layer, which contained the vector, was transferred to a new 1.5 ml tubes using a wide bore tip. (**N.B.** This step was frequently omitted and the extraction would start with a step (ii)).

   (ii) Then an equal volume of CI was added. The incubation was carried out for 2 min at room temperature. The sample was centrifuged at 13,000 rpm for 3 min. The top (aqueous) layer was transferred to a new 1.5 ml tube without disturbing the interface.

(iii) The precipitation with 2.5x volume of absolute ethanol and 1/10 volume of 2M NaCl was carried out at –20°C overnight, or at –64°C for 15-20 min.

(iv) The sample was centrifuged at 13,000 rpm for 15 – 20 min, The pellet was washed twice with 70% ethanol, dried and re-suspended in 0.1 x TE buffer.

The T-tailed vector was ready for insertion of a selected RAPD-PCR DNA fragment, i.e. for a ligation procedure.

Ligation:

The ligation reaction contained:

12 µl sd $H_2O$
2 µl 10x Ligase Buffer (contain ATP) (Promega, or NBL Gene Sciences)
2 µl T- tailed vector
2 µl RAPD-PCR insert (fragment)
2 µl T4 DNA ligase (3-4 units/ µl) (Promega, or NBL Gene Sciences)
20 µl – total reaction volume

The T4 DNA ligase (3 u/µl; Promega, Madison, WI, USA, or 4 u/µl; NBL Gene Sciences, Northumberland, UK) was used to ligate selected RAPD-PCR DNA fragments into *Sma* I cut/ T-tailed pUC18. Usually, a ligation reaction was incubated overnight at 12°C, but occasionally, variable temperature regimes were also applied (the ligation reaction was first incubated at 4°C overnight, then 5-10 hours at 12°C, then 30 min to 2 hours at room temperature, and to finish incubation at 4°C for a few hours) in order to achieve a better ligation effectiveness.

***Transformation***

*E. coli* XL2 - BlueScript (Stratagene) host bacterial cells were used for transformation. Commercially prepared competent cells were stored at -70°C until required for transformation, or competent cells were prepared from a culture. If preparation of competent cells was required, two protocols were applied in the study:

(1) modified Kushner method (from: Hammond 1995) for inducing competence and transformation by calcium chloride/rubidium chloride procedure, and
(2) modified calcium chloride protocol (Cohen *et al.* 1972).

**(1) <u>The Kushner modified method (from: Hammond 1995)</u>**

**(a)** <u>Production of competent cells</u>

1) A loopful of *E. coli* XL2 cells from a single colony was inoculated into 10 ml of LB broth. The cells grew at 37°C overnight in an incubator (gentle shaking).

2) 0.1 ml of the overnight culture was inoculated into 20 ml of LB broth in a 250 ml conical flask and shaken at 37°C for 2-3 hours, until the bacterial growth reached early to mid-log phase (measured by a spectrophotometer: $OD_{650} = 0.15$).

3) Then, 1.5 ml of the bacterial cell culture was transferred to 1.5 ml Eppendorf tubes and pelleted by spinning at 12,000 rpm in a microcentrifuge for 2 minutes. The supernatant was removed and the pellet gently re-suspended in 0.5 ml of ice-cold 10 mM MOPS, 10 mM RbCl pH 7.0. Cells were kept on ice for all subsequent steps unless otherwise stated.

4) The suspension was centrifuged, the supernatant removed and the pellet re-suspended in 0.5 ml of ice cold 100 mM MOPS, 50 mM $CaCl_2$, 10 mM RbCl, pH 6.5. The cells were then incubated at 0°C for 60 minutes.

5) The suspension was centrifuged for 2-5 minutes at 12,000 rpm, the supernatant decanted, and then the pellet re-suspended in 200 µl of ice cold 100 mM MOPS, 50 mM $CaCl_2$, 10 mM RbCl, pH 6.5.

At this stage, the cells were "competent" and they were immediately used for introduction of plasmid, or they were stored in glycerol (40 µl of 80 % sterile glycerol + 200 µl of cell suspension) at -70°C until required. When required, they were thawed on ice and immediately used for transformation.

**(b)** <u>Introduction of plasmid into competent cells</u>

1) 4 µl of the ligation reaction was pipetted directly into 100 µl of competent cell suspension and gently mixed. This mixture was incubated on ice for 30 minutes with occasional shaking.

2) A heat shock of 45°C for 30 s was applied.

3) The tubes were removed from the 45°C thermal block (or water bath) and 1 ml of warm LB broth was added. The mixture was incubated for 1 hour at 37°C.

4) Volumes of 50 µl, 100 µl, 150 µl, and 200 µl from each transformation were plated out on separate LB/X-gal/IPTG plates containing 50 µg/ml of ampicillin (Sigma).

5) Plates were placed in an incubator at 37°C for 24 – 48 hours.

**(2) <u>The modified method of Cohen *et al.* (1972) for inducing competence to host cells and transformation</u>**

1) 10 ml of LB broth was inoculated with 500 µl - 1000 µl of an overnight bacterial culture. The cells were incubated at 37°C and shaken until they reached exponential phase. Then, the cell culture was centrifuged for 10 min at 5,000 rpm at 4°C.

2) The supernatant was discarded and then the cell pellet was gently re-suspended in 5 ml of ice cold 50 mM $CaCl_2$. This cell suspension was placed on ice for 30 min

3) Centrifugation at 5,000 rpm for 10 min at 4°C was then applied. The supernatant was discarded. The cells were re-suspended in 1 ml of 50 mM $CaCl_2$ on ice. At this stage, the cells became "competent" and ready for transformation.

4) The cell suspension was dispensed in 200 µl aliquots into pre-chilled tubes. 10-50 µl of ligation mix (pUC18 with inserted RAPD-PCR fragment) was added to each aliquot. The mixture was placed on ice for 30 min. Positive (with pUC18) and blank (with water) controls were also carried out.

5) A heat shock of 10 min at 37°C was applied.

6) 1 ml of cold LB broth was added into each tube and cells were incubated at 37°C for 1 hour.

7) Volumes of 50 µl, 100 µl, 150 µl, and 200 µl from each transformation were spread on separate LB/X-gal/IPTG plates containing 50 µg/ml of ampicillin. Plates were placed in an incubator at 37°C for 24 – 48 hours.

Plasmid isolation

Selected white colonies were re-plated and grew into LB broth (containing ampicillin), assuming that they contain recombinant plasmids (plasmids with cloned RAPD-PCR fragment). Plasmid DNA preparations were carried out on a small-scale using Wizard Minipreps DNA Purification System (Promega, Madison, WI, USA) or Hybaid Recovery Quick Mini Spin Kit (Hybaid, Middlesex, UK), and on a large-scale using Wizard Maxipreps DNA Purification System (Promega, Madison, WI, USA). The procedures were performed according to the manufacturers' instructions.

***Screening methods:***

Determination of the presence and size of inserts in recombinant plasmids (plasmids isolated from white colonies) used:

(1) PCR screening method (occasionally), or

(2) Screening method that involves *EcoR* I/ *Hind* III plasmid digestion (a method most often applied in this study).

**(1) PCR screening method**

The pUC18-specific PCR primers applied in this study:

Primer 1 (forward), PUCFO: **5`-cag ggt ttt ccc agt cac gac-3`**

Primer 2 (reverse), PUCRE: **5`-tca cac agg aaa cag cta tga c-3`.**

(Hammond 1997, personal communication)

For PCR reactions, a 10x diluted DNA suspension of isolated recombinant plasmid was used as a PCR-DNA template. The PCR reaction mix contained: 0.5 µl template DNA; 1x *Taq* polymerase reaction buffer (Hybaid AGS*Gold Taq* polymerase buffer without Mg), 2.5 mM $MgCl_2$; 400 µM dNTP mix; 200 nM of each primers; 1 unit AGS*Gold Taq* polymerase (Hybaid), in a final reaction volume of 25 µl. The PCR reactions were carried out using the Hybaid Omn-E thermal cycler programmed as follows: initial denaturation at 94°C for 2 minutes; then 35 cycles of: 30 s at 94°C; 30 s at 54°C; 1 min 30 s at 72°C. The reaction products were analysed by agarose electrophoresis on a 1 % - 2.5 % agarose gel.

**(2) <u>Screening white colonies by applying *EcoR* I and *Hind* III restriction enzyimes</u>**

Plasmid DNA from white colonies was analysed for size and the presence of inserts by digestion with two restriction enzymes (*Eco*R I and *Hind* III). If the insertion (ligation) and recombination were successful, the RAPD-PCR fragment (insert) should be incorporated between these two restriction sites. For routine checking of the insert presence, the total volume of the digestion reaction was usually 20-30 µl. If the insert was required to be cut from the gel in order to be applied for DIG labelling and Southern blotting, or RFLP analysis – the digestion reaction was usually prepared in 80-100 µl of volume. The length of the incubation time (3-5 hours at 37°C) depended on the reaction volume used. Digestion reactions contained:

| | | |
|---|---|---|
| 15 µl plasmid DNA suspension | **OR** | 40 µl plasmid DNA suspension |
| 10 µl sterile distilled water | | 26 µl sterile distilled water |
| 3 µl multi-core buffer (Promega) | | 8 µl multi-core buffer (Promega) |
| 1 µl *Hind* III (12 U/µl; Promega) | | 3 µl *Hind* III (12 U/µl; Promega) |
| <u>1 µl *Eco*R I (12 U/µl; Promega)</u> | | <u>3 µl *Eco*R I (12 U/µl; Promega)</u> |
| **Total reaction volume: 30 µl** | | **Total reaction volume: 80 µl** |

The applied digestion conditions:

| | | |
|---|---|---|
| 1) 3 hours at 37°C | **OR** | 1) 5 hours at 37°C |
| 2) 15 minutes at 55°C | | 2) 20-30 minutes at 55°C |

For a routine check, 3 - 5 µl of digested recombinant plasmid and 2 µl of loading buffer were mixed and run on a 1% - 2.5 % agarose gel (the percentage of the agarose gel depended on the expected size of inserted RAPD-PCR fragment). If the insert (ligated RAPD-DNA fragment) was to be excised from the gel and used for another analysis (e.g. for Southern blotting or RFLP), 30 - 50 µl of the digested recombinant plasmid was loaded into large wells and run on a 1.5% - 2 % agarose gel. A desired DNA fragment was excised from the gel and purified by Hybaid Recovery Gel Extraction Kit.

### *2.2.3.5. RFLP experiments*

Restriction Fragment Length Polymorphism analysis (RFLP) was applied to:

- selected cloned inserts (cloned RAPD-PCR DNA fragments),
- uncut recombinant plasmids
- RAPD-PCR fragments excised directly from the gel slice after RAPD-PCRs,
- crude PCR products (RAPD-PCR amplifications that generated only one visible PCR product)
- total genomic DNA extracted from a formalin tissue sample of *Nezumia*.

Boehringer Mannheim restriction nucleases: ***Alu* I** (AG*CT), ***Ava* I** (C*YCGRG), ***Ava* II** (G*GWCC), ***Bgl* I** (GCCNNNN*NGGC), ***Bcl* I** (T*GATCA), ***Cla* I** (AT*CGAT), ***Dde* I** (C*TNAG) ***Nae* I** (GCC*GGC), ***Nde* I** (CA*TATG), ***Sac* I** (GAGCT*C), ***Sma* I** (CCC*GGG), ***Stu* I** (AGG*CCT), ***Xho* I** (C*TCGAG) were used.

Digestion reactions were carried out in a total volume of 15 - 20 µl using 5 µl of cloned insert (recombinant plasmid, or PCR amplification), 1/10 volume of the appropriate buffer, and 2 µl of a restriction enzyme. Length of digestion was at least 3 hours at the appropriate temperature (according to the manufacturer's information sheet) for each restriction enzyme. Digestions were stopped by applying a temperature of 60°C for 10-15 min

### *2.2.3.6. Sequencing*

Sequencing was performed on:

- recombinant plasmids, i.e. cloned RAPD-PCR fragments, and
- PCR products generated with *Nezumia*-specific (RAPD-derived) primers and mitochondrial 16S primers (direct sequencing).

Automated four colour sequencing employing cycle-sequencing dye-terminator chemistry was performed on ABI model DNA sequencers (Perkin-Elmer Corp.) by Cambridge Bioscience Limited (sequencing of cloned RAPD-DNA fragments), MWG-BIOTECH AG and the Natural History Museum, London (direct sequencing of PCR products). Short read <650/single strand sequencing with primers M13(-21) was performed for sequencing of

cloned RAPD fragments. Direct sequencing of PCR products (both strands) was performed with the same primers that were used for PCR amplifications.

Sample purification, cycle sequencing and sequencing of cloned RAPD-DNA fragments and direct sequencing of PCR products were completely carried out by the service staff in Cambridge Bioscience Limited, UK (suspensions of the recombinant plasmids that contained inserted RAPD-DNA fragments were sent to them) and MWG-Biotech (amplified PCR products – the vials with PCR reactions were sent out), but for direct sequencing performed in the NHM laboratory, the preparation of samples for sequencing was carried out to the point of purifying sequencing reaction product (the rest of the sequencing procedure was carried out by the staff of the NHM Sequencing Facility Unit). Re-amplified PCR product was purified using QIAquick PCR Purification Kit (Qiagen) following the supplier's instructions. Purified PCR product was used in a cycle sequencing reaction using ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit (PE Applied Biosystems). A cycle sequencing was performed in 10 µl reaction containing: 4 µl Dye mix, 1 µl primer (1.6 pmol), 4 µl sample - purified PCR-DNA product (~40 ng), and 1 µl sd$H_2O$. A thermal cycler, GeneAmp PCR System 2700, was programmed: 1 cycle at 96°C for 5 min; 25 cycles of 96°C for 20 s, 50°C for 10 s, 60°C for 4 min; and a holding step at 4°C. Extension products were purified by using ethanol/sodium acetate precipitation method (in microcentrifuge tubes) following the supplier's instructions (ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit; PE Applied Biosystems), except that half of the recommended amounts of sodium acetate (1 µl instead of 2 µl) and ethanol (25 µl instead of 50 µl) were used. The samples (dry pellets of the purified extension reaction) were stored at -20°C in the Sequencing Facility Section of the Natural History Museum London where sequencing was performed using the ABI Prism 377 automated sequencer.

Sequences were analysed by the Sequencher package sequence analysis program (Macintosh version 4.1; Gene Codes Corp., Ann Arbor, MI, USA) – sequence editing, sequence alignments and contig assembling. Sequences were submitted to the GenBank using the Sequin program (Macintosh version) downloaded from the NCBI World Wide Web site (http://www.ncbi.nlm.nih.gov).

## 2.3. ***Statistical analysis***

Tests of PCR success were checked on standard agarose gels stained with ethidium bromide. A weighting regime was applied as follows: any clear band/s (bright/strong or weak/clear) was considered as "positive (+)" PCR, non-amplified products were considered as "negative (-)", whereas very weak/unclear bands were treated as possible artifacts ("positive/negative (+/-)" PCRs) and were excluded from statistical calculations. Statistical analysis of data was performed in the Biometric Group of the Natural History Museum, London (with the assistance of Clive Moncrieff) by using GenStat 10.2 software package (2007; Lawes Agricultural Trust – Rothamsted Experimental Station, U.K.). $\chi^2$ analyses were undertaken to compare PCR success rate (selected mitochondrial and RAPD PCRs) related to DNA extraction method applied to formalin-fixed, Steedman's preserved specimens of *Nezumia*, whereas t-tests were used to compare the number of RAPD bands in conjunction with an extraction protocol.

# Chapter 3. RESULTS AND DISCUSSION

Research on deep-sea organisms and use of museum specimens (or preserved tissue samples from other collections) for molecular investigations are two aspects of biological science of increasing interest. The main goal of this project was to explore the feasibility of using museum-preserved specimens for DNA studies. *Nezumia* was used as a test material and model approach for this kind of study. The type of preservation used on investigated *Nezumia* specimens (fixation with unbuffered formalin and preservation in Steedman's solution for prolonged periods) is of particular interest because there is no published information for molecular studies on such preserved specimens (or they were unsuccessful; see Carter 2003) despite the fact that many of museum fish and amphibian specimens have been preserved and stored in Steedman's solution. *Nezumia* is also an excellent model for investigating the feasibility of molecular investigations on organisms that have unstudied genomes and are available mostly as preserved specimens. This study did not aim to construct/re-construct phylogenetic relationship within *Nezumia* species and macrourid fishes, but to investigate and develop possible approaches in obtaining molecular information from fluid-preserved fish collections and taxa with unstudied genomes. The intention also was to contribute to modern taxonomy (if possible) in building up the "genetic inventory" of *Nezumia* species by using exclusively museum specimens of two investigated species.

Throughout the text the term "archival DNA (arDNA)" will be used for DNA extracted from preserved specimens, but discussion about this new term and justification for introducing this term will be discussed in the Section 3.5 (p 228) of this thesis.

## *3.1. Extraction of DNA and the effectiveness of PCR with "arDNA"*

This section contains two parts of the investigations conducted in the project:

(1) Testing and developing the most suitable method for formalin-arDNA extraction including investigations of different effects on quality and quantity of extracted DNA if different protocols and modifications are applied, as well as the effect on PCR performance;

**(2)** Development of appropriate approaches for RAPD-PCR investigation with formalin-arDNA; that is, the use of RAPDs for testing the efficiency of DNA extractions/amplifications.

### 3.1.1. DNA retrieval from preserved specimens and DNA extraction protocols

The recovery of DNA from formalin-fixed specimens is challenging and an optimal method for DNA retrieval from such specimens (archival collections) is not yet described. The exact effects of different pre-extraction treatments of preserved tissue (washing/ heating/ freezing/ drying of preserved tissue prior to the DNA extraction) and DNA extraction methods on DNA recovery are still largely unknown (Tang 2006).

#### *3.1.1.1. Developing and optimising a protocol for formalin-arDNA extractions*

As was emphasised previously, the important part of this project involved the optimisation and development of a reliable protocol for extraction of formalin-arDNA suitable for PCRs and other molecular analysis. None of the initially tested protocols (protocols (B), (C) and (D)) for extracting of DNA from formalin-fixed, Steedman's-preserved specimens gave completely satisfactory results. This was the reason for developing and optimising a new DNA extraction protocol (protocol (A)) suitable for formalin-Steedman-arDNA extractions. The protocol (A) developed and optimised in this study was the result of gathering information obtained by testing the above-mentioned three protocols, including pre-extraction treatments of preserved tissues. Problems encountered with existing DNA extraction protocols, and how "good parts" of these protocols were tested in order to develop and optimise protocol (A), will be described. Guanidinium-based protocols were first tested (protocols (B) and (D)), and then a phenol-based protocol (C).

### 3.1.1.1.1. Tests and problems encountered with guanidinium-based protocols (B) and (D)

The GES based protocol of Hammond *et al*. (1996), **protocol (B)**, provided arDNA extracts with relatively good PCR performances (Fig. 8). With these extracts, a certain level of RAPD-PCR reproducibility was obtained (Fig. 8, gel (b)-Lanes: 1-4). However, some arDNA extracts produced by this protocol were not suitable for all PCR amplifications, especially not for mitochondrial PCRs (see support material on the CD for more information). PCR inhibition was recorded with some arDNA extracts in combination with particular RAPD primers (Fig. 8, gel (a)), but most of these arDNAs became PCR amplifiable by diluting arDNA extracts (Fig. 8). This indicated that this protocol could produce arDNA extracts with sufficient concentration and quality of DNA to generate successful RAPD-PCR amplifications and reasonably reproducible band patterns in PCR experiments with some of arDNA extracts (e.g. Fig. 8, gel (b)-Lane: 1-4).

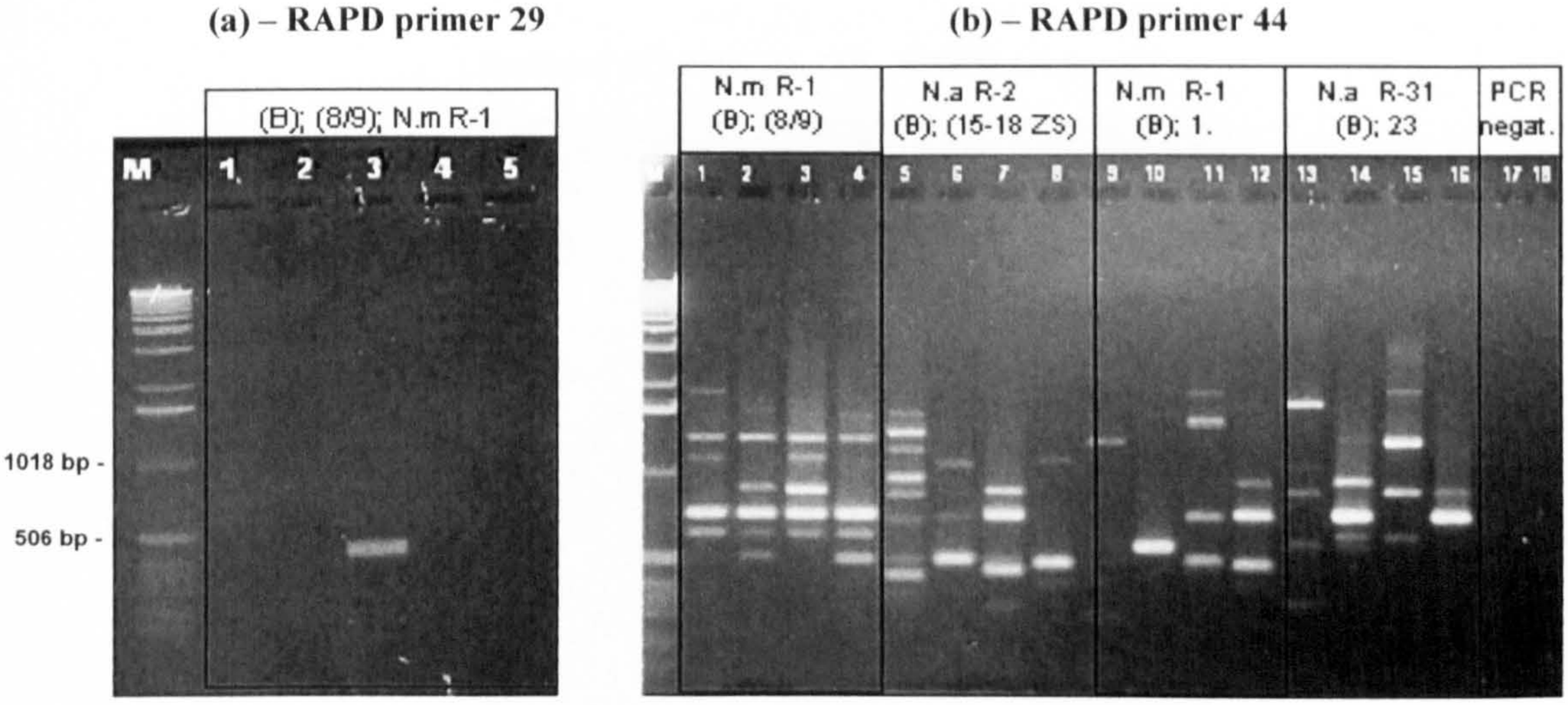



**Fig. 8 – RAPD-PCR amplifications** performed with four different arDNA extracts produced by protocol (B).
**(a)** – Possible PCR inhibition with undiluted and 5x-diluted (5x-D) arDNA suspension (Lanes: 1 and 2) - RAPD primer 29. Lack of PCR amplifications with 50x- and 100x-diluted (100x-D) arDNA samples were probably due to too low DNA-template concentrations (Lanes: 4 and 5).
**(b)** – RAPD-PCRs performed in duplicates with two different PCR-DNA template concentrations (all arDNA extracts were performed with 5x-dilutions of the original arDNA suspension). The reproducibility of RAPDs was variable and depended on individual arDNA extracts and fish individuals (RAPD-PCRs performed with RAPD primer 44).
"M" indicates size marker fragments - 1Kb DNA Ladder (Gibco).
**N.a** – *Nezumia aequalis*
**N.m** – *Nezumia micronychodon*
**R** – the identifier of the fish individuals (e.g., R-31 is fish no. 31, R-2 is fish no. 2., etc.)
**(B)** – arDNA extracts produced by protocol (B)

However, this strategy of diluting arDNA extracts with $sdH_2O$ was not successful with all tested arDNA extracts. Generally, the concentrations of DNA in arDNA extracts produced with protocol (B) were relatively low (usually 1-2 μg/ml; Appendix 1) and a few arDNA extracts were so low that it was not possible to measure DNA concentration by applying the Saran method (Sambrook *et al.* 1989). In addition, a certain number of PCR amplifications did not provide reproducible RAPD-PCR fragments (reproducible banding patterns) even in duplicates (e.g., Fig. 8, gel (b)-Lanes: 5-16). All this urged improvements and modification of the protocol (B), i.e. the testing of other DNA extraction protocols.

The assumption that guanidinium was suitable for extractions from formalin-fixed, Steedman's-preserved tissue samples had been further supported by PCR results obtained with arDNA extracts produced by **protocol (D)** - the Genosys Isolator reagent (a reagent that also contain guanidinium according to the supplier's information). This protocol produced arDNA extracts that gave some good PCR amplifications and RAPD-PCR results (Fig. 9 with diluted arDNA extracts: Lanes 3-5).

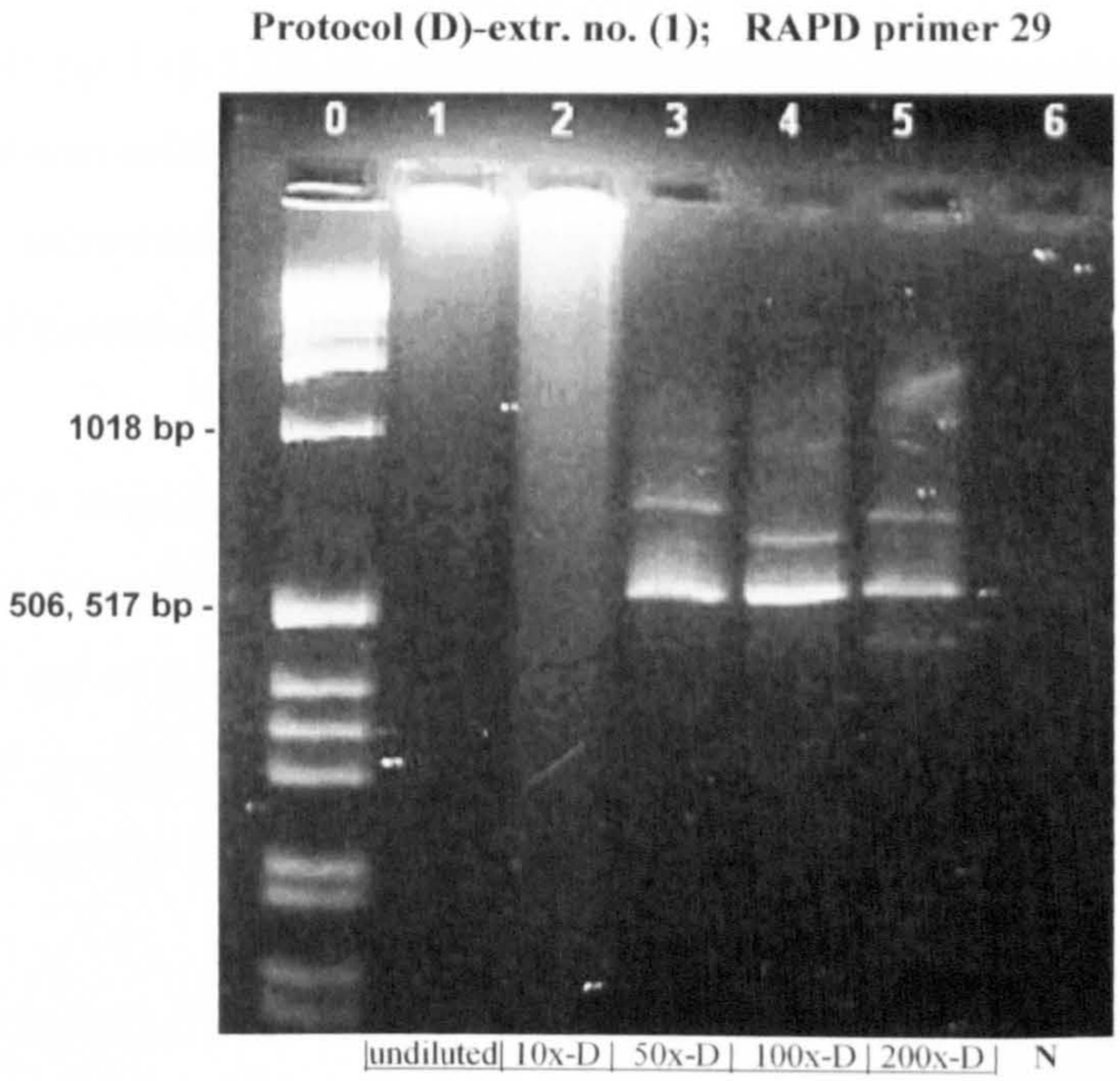



**Figure 9** – The arDNA extract no. (1), produced by modified protocol (D), exhibited the inhibition of the RAPD-PCR if undiluted and 10x-diluted arDNA extracts were applied (Lanes: 1 and 2). It was possible to reduce (eliminate) the PCR inhibitory effect by diluting arDNA extract with sterile water:
50x-diluted (50x-D) (Lane 3),
100x-diluted (100x-D) (Lane 4),
200x-diluted (200x-D) (Lane 5).
RAPD-PCR reactions were prepared from the same PCR master mix (mixture of the RAPD primer, *Taq* Buffer, *Taq* polymerase, dNTPs, $MgCl_2$ and $sdH_2O$) and they were run simultaneously in the same PCR thermal cycler.

The concentrations of DNA in the extracts of cod (frozen tissue that was used as a control) were high (usually over 20 μg/ml) by applying protocols (B) and (D) – Appendix 1 (rows VII and VIII), suggesting a high efficiency of these two protocols in extracting DNA from "good" samples. However, DNA concentrations of archival DNA extracts produced by these protocols were low, and PCR performances were inconsistent. Protocol (D) in particular, original or modified, was not considered for adoption as a main protocol throughout the project because of the inconsistency in producing PCR amplifiable arDNA (see support material on the CD for more information), but it helped to gain knowledge of the factors that are important for isolating arDNA from formalin-fixed fish samples.

#### 3.1.1.1.2. Tests and problems encountered with phenol-based protocol (C)

The Shedlock *et al.* (1997) phenol-based protocol (protocol (C)) provided a better yield of formalin-arDNA (Appendices 2 and 3; Table 6, p 140) than protocols (B) and (D), but produced arDNA extracts which exhibited a very low level of PCR usability (Figs. 10 and 11; Tables 3 and 4; pp 131-133). A strong PCR inhibition was recorded with most of the tested formalin-Steedman-arDNA extracts produced by protocol (C). Almost all RAPD-PCR amplifications were completely inhibited (Fig. 10), as well as mitochondrial PCRs (Fig. 11) if the unmodified protocol of Shedlock *et al.* (1997) was applied.

Most of the PCR amplifications run on the 1% agarose gel exhibited very strong background in the lanes (Figs. 10(a) and 11). The explanation and importance of the strong background on the agarose gel of the PCRs visualised by UV were unclear, but it was usually associated with poor PCR performances and poor quality of DNA template (arDNA extract). Similar observations were also presented by Chakraborty *et al.* (2006).

In general, protocol (C) did not prove to be a good DNA extraction protocol for the fish specimens investigated because of strong PCR inhibition. Some improvements in PCR amplifications were achieved by diluting arDNA extracts with sterile water, or by applying the Promega Wizard Clean Up Kit instead of phenol and then diluting these arDNA extracts (Fig. 10 – gel (b)-Lanes: 1-6 and 13-18). However, neither of these modifications made to the protocol (C) significantly improved the protocol to the extent that it would be selected as a reliable DNA extraction protocol in the project.

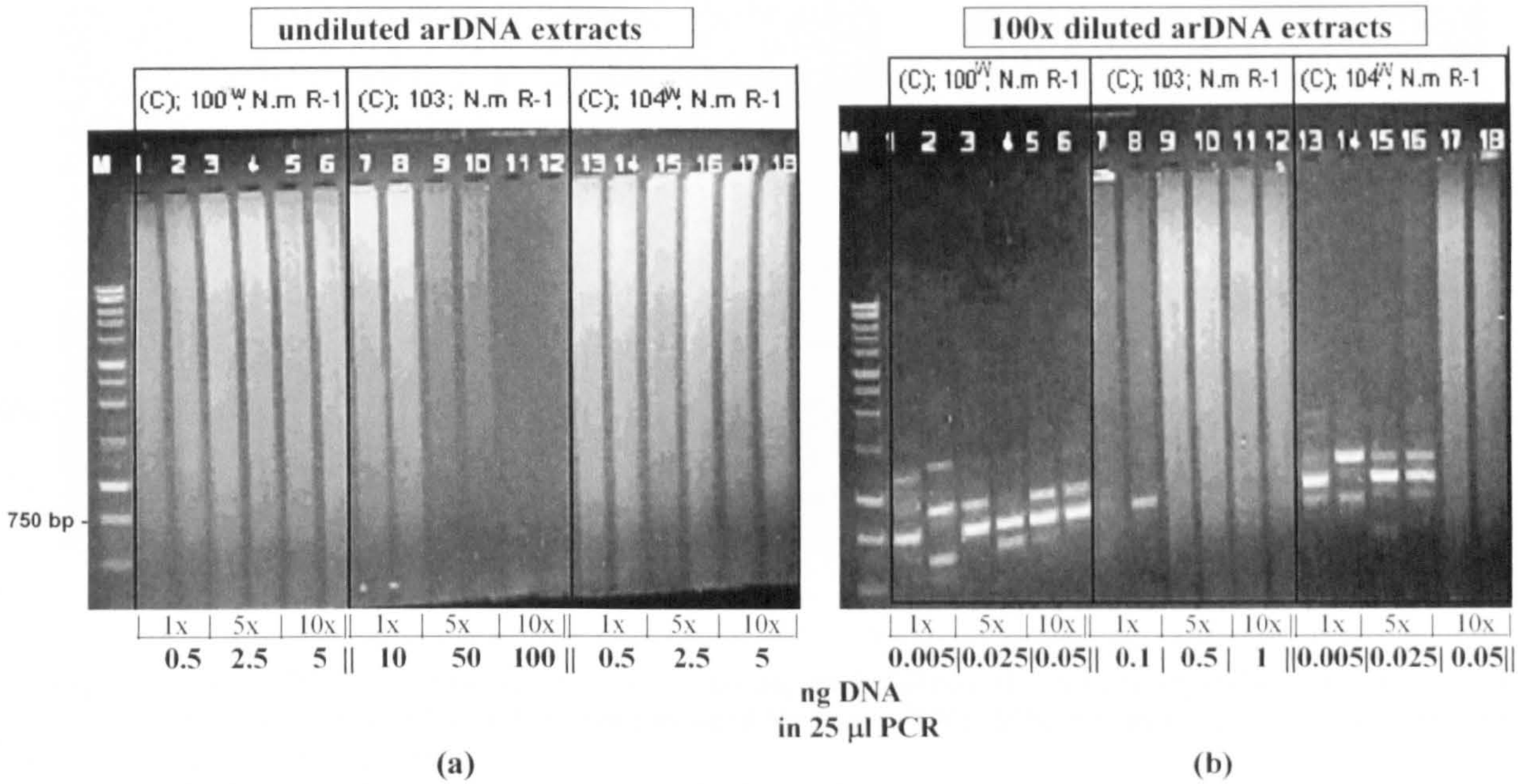
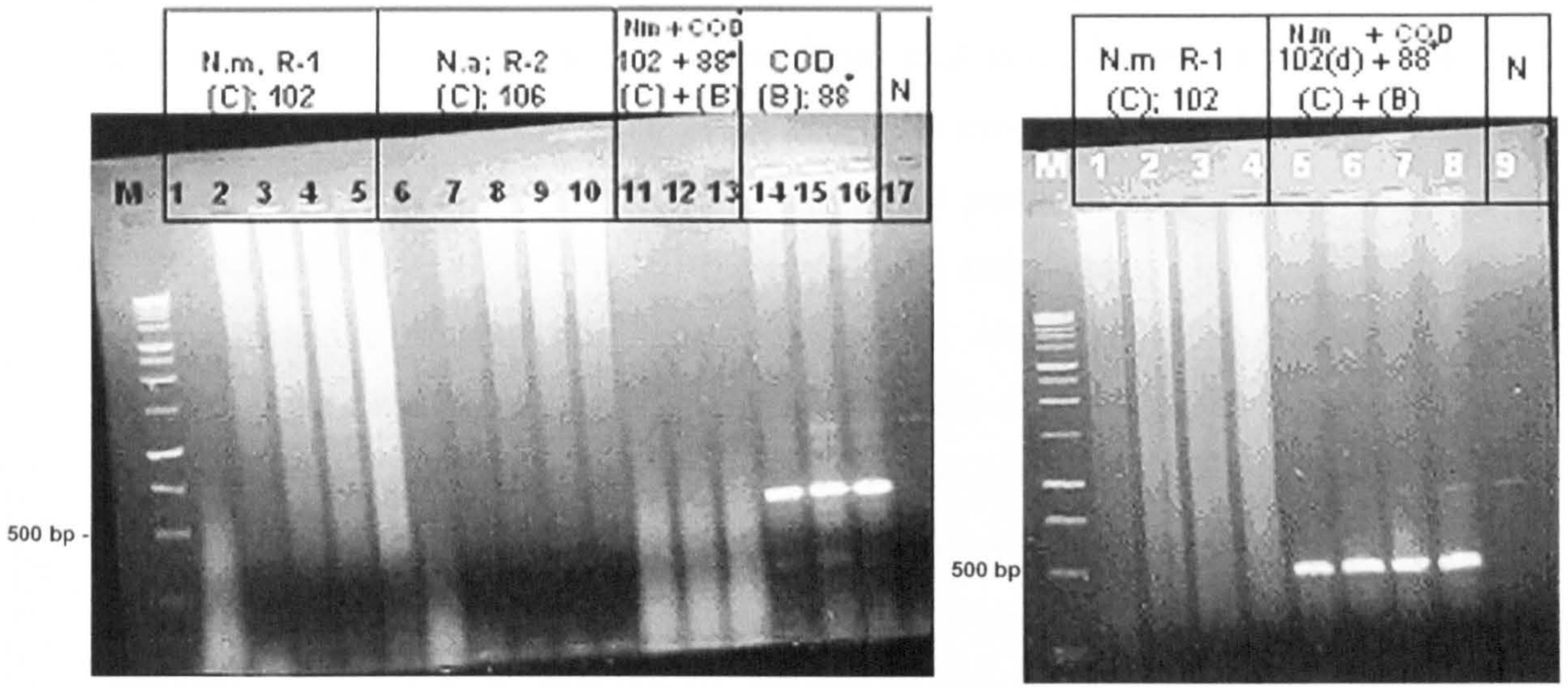



**Fig. 10** – RAPD-PCR amplifications with three different arDNA extracts produced by original (Lanes: 7-12 in both gels) and modified (Lanes: 1-6 and 13-18 in both gels) protocol (C). Protocol (C) was modified by using Promega Wizard Clean Up Kit instead of phenol. PCR amplifications were performed in duplicates of three different PCR-DNA template concentrations (relative concentrations: 1x, 5x and 10x). **Gel (a)** represents RAPD-PCR amplifications with undiluted arDNA extracts, whereas **gel (b)** represents RAPD-PCR amplifications with 100x dilutions of arDNA extracts that were used for PCRs presented in gel (a) – arDNA extracts nos.: $100^{W}$, 103, and $104^{W}$.
**"M"** indicates size marker fragments - 1Kb DNA Ladder (Promega).
**N.m** – *Nezumia micronychodon*
**R** - the identifier of the fish individuals (R-1 is fish no. 1)
**(C)** – arDNA extracts produced by protocol (C).

The presence of diffusible PCR inhibitors in arDNA extracts is supported by evidence presented in Figure 11. The experiment presented on the gel (a)-Lanes: 11-13 clearly exhibits the PCR inhibition caused by diffusible PCR inhibitors from arDNA extract of *Nezumia* if two DNA extracts are mixed (*Nezumia*-arDNA extract no. 102 produced by protocol (C) and control cod DNA extract no. $88^{+}$ produced by protocol (B)). DNA of cod alone generated PCR products (Fig. 11, gel (a)-Lanes: 14-16) suggesting that the quality and concentration of DNA of the cod were sufficient for successful PCR amplifications. However, successful PCR amplifications were prevented by diffusible PCR inhibitors from arDNA extracts of *Nezumia* if these two DNA extracts were mixed (Fig. 11, gel (a)-Lanes: 11-13).




**(a)** **(b)**

**Fig. 11** – Strong PCR inhibition was proved by mixing control DNA of frozen cod (produced by protocol (B)) and formalin-arDNA of *Nezumia* produced by protocol (C). Mitochondrial PCR amplifications of the COIII gene were tested.

**Undiluted arDNA extracts**: Lanes 1 and 6 on gel (a).

**Diluted arDNA extracts:** 10x [Lanes: 2 and 7 on gel (a), and Lane 1 on gel (b)],
50x [Lanes: 3 and 8 on gel (a)],
100x [Lanes: 4 and 9 on gel (a), and Lane 2 on gel (b)],
200x [Lanes: 5 and 10 on gel (a), and Lane 3 on gel (b)],
1000x [Lane 4 on gel (b)].

The control DNA extract (no. $88^{+}$) from the fresh/frozen cod specimen generated successful mitochondrial amplifications alone (undiluted DNA extract; Lanes 14-16 on gel (a)).

**The presence of PCR inhibitors in formalin-arDNA extracts produced by protocol (C)** is demonstrated on gel (a) – Lanes: 11-13 by mixing undiluted formalin-arDNA extract no. 102 of *Nezumia micronychodon* produced by protocol (C) and cod control DNA extract (no. $88^{+}$) produced by protocol (B). The PCR products were generated only if *Nezumia* arDNA extract was diluted 10x-1000x and then mixed with undiluted cod control extract no. $88^{+}$ (Lanes: 5-8 on the gel (b)).

**"M"** indicates size marker fragments - 1Kb DNA Ladder (Promega).

**N.a** – *Nezumia aequalis*

**N.m** – *Nezumia micronychodon*

**R** - the identifier of the fish individuals (R-1 is fish no. 1)

**(C)** – arDNA extracts produced by protocol (C).

**(B)** – DNA extracts produced by protocol (B).

By diluting arDNA of *Nezumia* 10x-1000x times, diffusible PCR inhibitors were obviously diluted sufficiently to allow the successful PCR amplifications of cod when mixed with the diluted arDNA extract of *Nezumia* (Figure 11; gel (b)-Lanes: 5-8). However, dilutions of *Nezumia*-arDNA no. 102 alone did not produce successful PCR amplification (Figure 11, gel (a)-Lanes: 2-5 and gel (b)-Lanes: 1-4) although the applied concentrations of *Nezumia* arDNA (up to 5 ng DNA) should be sufficient for generating PCR products.

Generally, PCR performance was much worse with arDNA extracts produced by phenol protocol (C) than by guanidinium-based DNA extraction protocols (B) and (D) ($p<0.001$ related to the PCR success rate; $\chi^2$ values from 13.8 to 42.4 - see Tables 3 and 4; pp 131-133), despite higher DNA concentrations in arDNA extracts produced by protocol (C) – see Table 6 (p 138). Chemicals applied in the extraction protocol (C) are known to act as strong PCR inhibitors (e.g., SDS, phenol; Qiagen 2002), and there is also some information on possible negative interactions (from the aspect of recovering DNA from formalin-fixed specimens) between phenol and formaldehyde under acidic and/or alkali conditions (Fig. 12). During DNA extractions with the phenol-based protocol, yellowish, gelatinous bits in the pellet were frequently noticed after DNA precipitation. These bits were mostly insoluble in water and/or TE buffer unless DNA samples were warmed. Even after heating, the precipitate would not be completely dissolved. This might indicate specific (unknown) interactions between chemicals used for phenol-based DNA extractions and chemicals that were used for the preservation of the specimens.

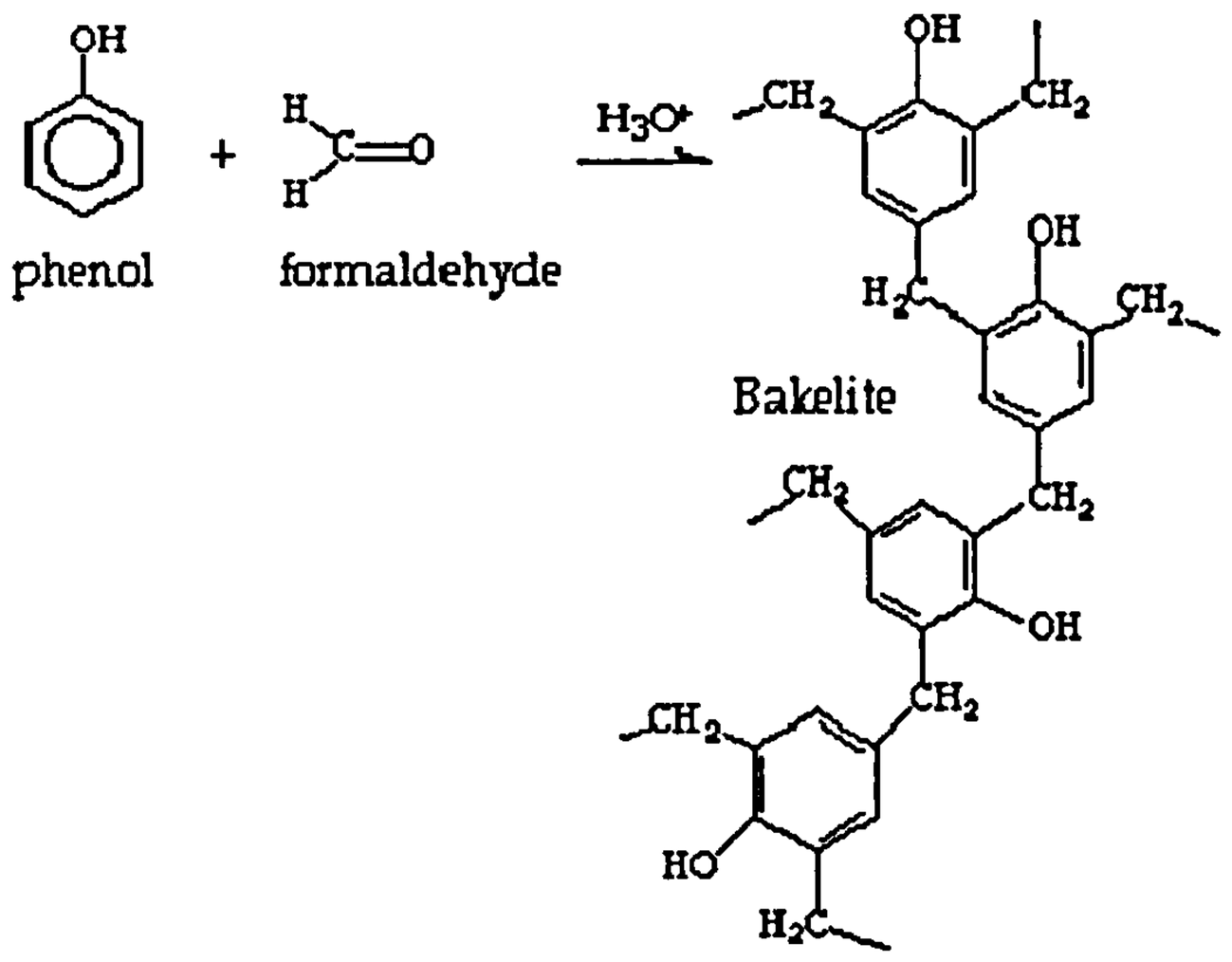



**Figure 12** - Phenols condense with formaldehyde under acidic or basic condition; polymerisation gives a network of phenol rings held together by methylene groups at ortho and para positions. (From Volland Walt: scidid.bcc.ctc.edu/wv/form-0.htm)

The possibility of obtaining a better yield of arDNA, as well as the possibility to make some small improvements in PCR performances by diluting these arDNA extracts, were good enough reasons to try some further modifications on protocol (C). First, additional purification of the DNA extracts with absolute ethanol was tested, or with absolute ethanol + 2 M NaCl, in combination with an increased number of pellet rinses with 70% ethanol. Small but insignificant improvements were observed. Next, reducing the amount of phenol

used was attempted. Steps **3** - (a) and (b) from the original protocol (C) were modified by reducing the use of phenol from three times to one, or by omitting (a) and (b) steps completely. Afterwards, instead of applying step **3** from the protocol (C), the GES reagent was used from the protocol (B), or the Promega Wizard Clean Up Kit. Applying these modifications (especially applying the GES reagent instead of phenol), the improvements were notable.

#### 3.1.1.1.3. Development and optimisation of a new protocol (A)

The results of investigations on the previously described three protocols aided development of a new protocol (A); that is, these helped in recognising "good and important parts" for successful formalin-Steedman's-arDNA extractions. Protocol (A) was mainly developed by combining and optimising the "parts" of protocols (B) and (C). A full and detailed description of this developed protocol is provided in Chapter 2: General material and methods (p 75).

The 1xGTE washing buffer (Shedlock *et al.* 1997) from protocol (C) was incorporated in protocol (A) as a part of pre-extraction treatment for formalin/Steedman's preserved tissues of *Nezumia*. This was based on claims that this buffer with glycine acts as a binding agent for excess formalin (Shedlock *et al.* 1997) which was supported by some tests and good results in this study. Interchangeable washing of a preserved tissue sample with 1xGTE buffer and $sdH_2O$ seems to be an even more effective pre-washing treatment (see Fig. 15; p 124). Also, it seems to be important that $sdH_2O$ is used as the last wash in this interchangeable washing treatment (see support material on the CD). This might be related to the fact that water is a good solvent and, by applying $sdH_2O$ as the last wash, it might aid in dissolving and washing out chemicals from tissue - chemicals which are components of 1xGTE buffer (such as glycine for which has been recently reported that can seriously decrease the production of amplifiable DNA if present in digestion buffer [see Gilbert *et al.* 2007b]), chemicals from preserved tissue, and/or chemicals that were dissolved by 1xGTE buffer, but not sufficiently removed from the tissue by this buffer. Others have also found it beneficial to use 1xGTE buffer for washing formalin-fixed tissue prior to DNA extraction (e.g. Kearney and Stuart 2004; Hasbun *et al.* 2005).

Fine grinding of the tissue using alumina powder and dry ice proved to be good for mechanical disruption of tissue. Combination of fast freezing by application of dry ice led to fragility of tissue samples, while grinding with alumina powder made a mechanical breakage of tissue more certain. Both applications (dry ice and grinding with alumina powder) also aided in drying a tissue sample.

The GES reagent from the protocol (B) was good for extraction of arDNA from preserved tissue, but it needed something additional for releasing DNA from the cells and tissue. Proteinase K proved to be useful for protein digestion and helped in releasing DNA from preserved tissue. The Shedlock *et al.* (1997) extraction (incubation; digestion) buffer with added DTT (dithiothreitol) as an antioxidant and enzyme stabilisor (information from the manufacture's booklet), proved to be effective (see Fig 10(b); p 116) and it was incorporated into protocol (A). Schander and Halanych (2003) are of the opinion that DTT is capable of breaking protein cross-linkages, i.e. DTT breaks the disulfide bonds between peptide sequences, whereas proteinase K cleaves peptide bonds at regular intervals (McNevin *et al.* 2005).

The Hammond *et al.* (1996) GES reagent proved to be useful in extracting DNA from formalin-fixed specimens (see Fig 8; p 113) and it was incorporated into a new protocol (A). This is probably because of the efficiency of guanidinium in denaturing proteins and dissolving biochemicals other than nucleic acids (Brown 2001; Rohland *et al.* 2004), as well as the possibility of breaking certain chemical cross-links (Rohland *et al.* 2004). The use of 7.5 M ammonium acetate (AmAc) and chloroform reagent (chloroform : pentanol, 24:1) from the Hammond *et al.* (1996) protocol (protocol (B)) were also incorporated into a new protocol (A), as well as the usage of isopropanol for DNA precipitation.

Isopropanol proved to be better than ethanol for precipitation of formalin-arDNA, because it produced arDNA extracts with better yields of DNA and with less PCR inhibitory effects in some tested arDNA extracts. This was the reason why isopropanol (not absolute ethanol) was incorporated into protocol (A) for DNA precipitation. DNA precipitation was tested with absolute ethanol (with and without NaCl, or ammonium acetate), but most arDNA extractions needed to be diluted in order to be PCR amplifiable (see support material on the CD). This might be related to the fact that ethanol precipitation of DNA can result in a significant loss of low-molecular-weight DNA (Mulligan 2005) and inability/inefficiency in removing PCR inhibitors (Micheli *et al.* 1994; Montiel *et al.* 1997). However, isopropanol

has generously been reported as being good (better than ethanol) for removing PCR inhibitors from DNA extracts (Hanni *et al.* 1995; Rohland *et al.* 2004). Extended washing of the DNA pellet (4-5 times) by 70% aqueous ethanol seems to be important in reducing the PCR inhibition, so this was incorporated as a part of protocol (A).

The protocol (A) provided a good DNA yield for preserved fish specimens (Appendices 4, 5; Table 6) and PCR performances were also good for the majority of arDNA extracts (e.g. Fig. 13; Tables 3-5, pp 131-134).

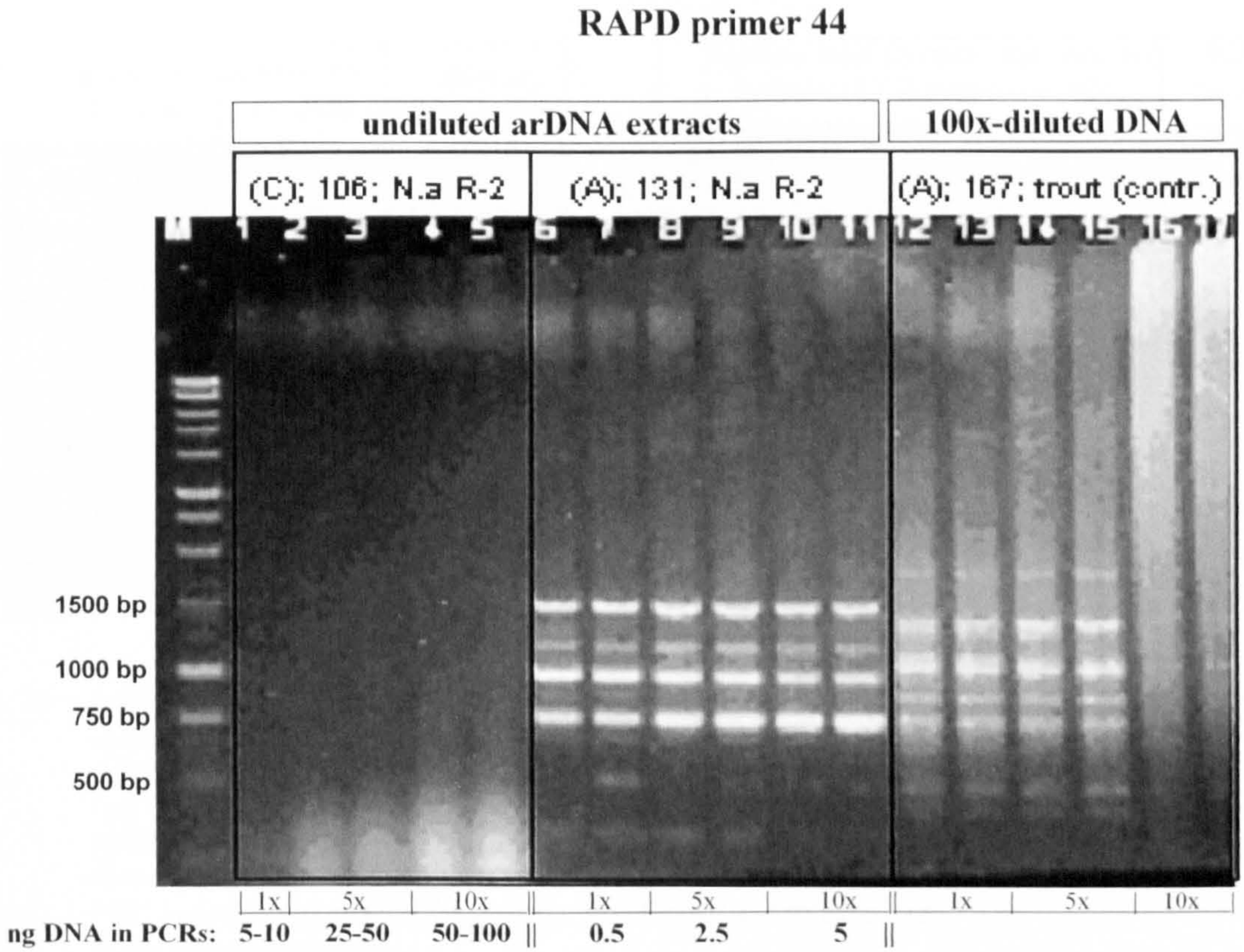



**Fig. 13** – RAPD-PCR amplifications (primer 44) with arDNA extracts produced by protocol (C) (Lanes: 1-5) and protocol (A) (Lanes: 6-11). Three different concentrations of PCR-DNA templates were applied: **1x**, **5x**, **10x** (relative concentrations of PCR-DNA templates). Control fresh tissue of rainbow trout was subjected to the complete procedure of washing and drying tissue for three days at room temperature before the DNA extraction procedure of protocol (A) was applied. This is not recommended for fresh/frozen tissue samples and this might be one of the reasons that the generated RAPD-PCR fragments were of a low intensity (Lanes: 12-17). The difference in RAPD-PCR profiles generated by primer 44 is notable between *Nezumia* and rainbow trout.
**"M"** indicates size marker fragments - 1Kb DNA Ladder (Promega).
**N.a** – *Nezumia aequalis*
**N.m** – *Nezumia micronychodon*
**R -** the identifier of the fish individuals (R-2 is fish individual no. 2)
**(C)** – arDNA extracts produced by protocol (C).
**(A)** – DNA extracts produced by protocol (A).

For the majority of arDNA extracts, there was no need to dilute them in order to dilute diffusible inhibitors and reduce PCR inhibitory effect using protocol (A). If a PCR amplification was unsuccessful with arDNA extract produced by protocol (A), the reason usually was not due to the presence of diffusible PCR inhibitors in arDNA extracts (Fig. 14), but related to other problems (may be because of extensive damage of targeted part of genome and the presence of cross-linkages [i.e. non-diffusible PCR inhibitors], insufficient amount of targeted DNA sequence in a PCR-DNA template, inadequate primers, and so on). This was proved by mixing formalin-arDNA of *Nezumia* and control DNA of rainbow trout (fresh/frozen tissue) (Fig. 14).

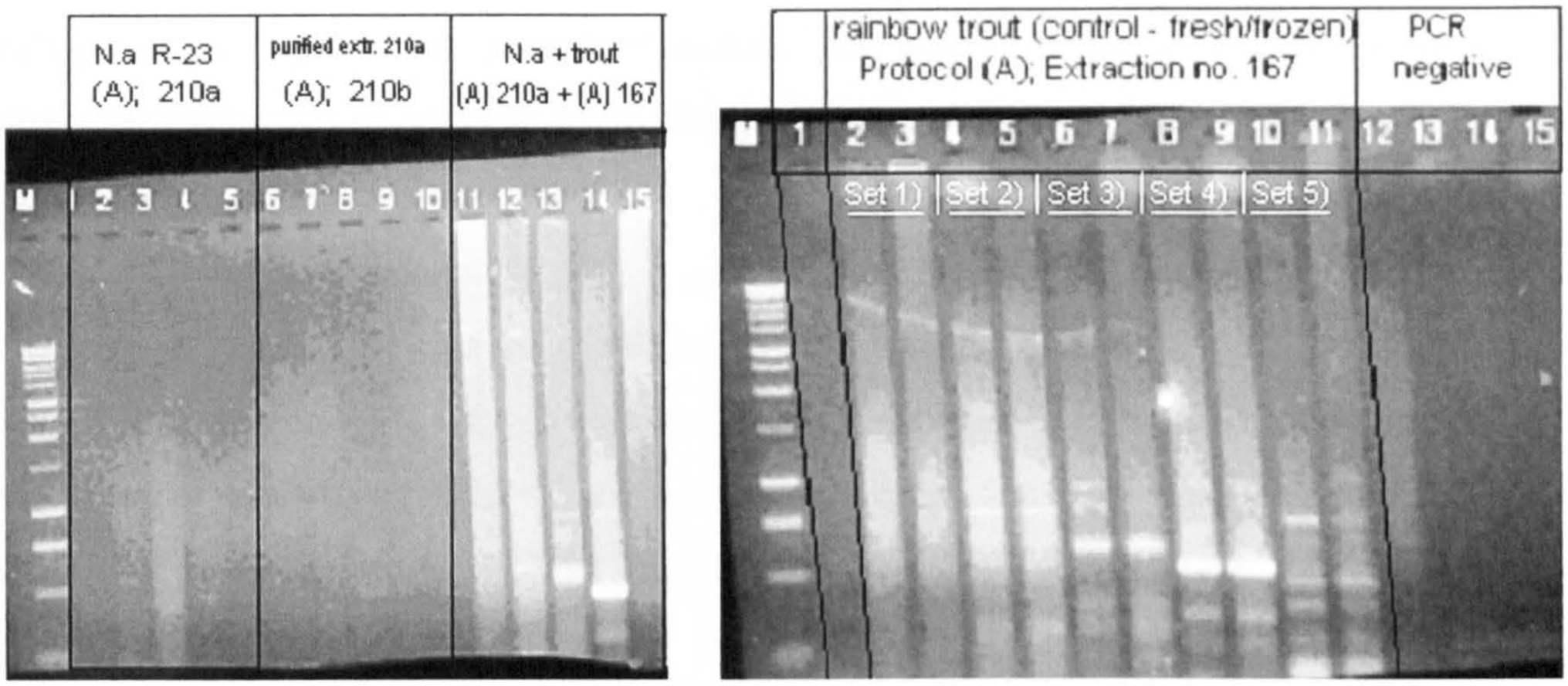



**Fig. 14** – Inability to amplify the regions of mitochondrial DNA in arDNA extracts produced by protocol (A) (Lanes: 1-10 on gel (a)) were not caused by "diffusible" PCR inhibitors. This was demonstrated by mixing control DNA of fresh/frozen rainbow trout and arDNA of formalin-fixed *Nezumia aequalis* (gel (a)-Lanes: 11-15). The same mitochondrial regions were PCR amplified in mixed DNA extracts (arDNA of *Nezumia* + DNA of rainbow trout) (Lanes: 11-15 on gel (a)) and in DNA extract of trout alone in duplicate PCR amplifications (Lanes: 2-11 on gel (b))
**"M"** indicates size marker fragments - 1Kb DNA Ladder (Promega).
**N.a** – *Nezumia aequalis*
**R** - the identifier of the fish individuals (R-23 is fish individual no. 23)
**(A)** – DNA extracts produced by protocol (A).

Generally, protocol (A) produced formalin-Steedman-arDNA extracts with better yield and PCR usability than previously tested protocols (B)-(D). On average, DNA concentrations were 6.70 μg/ml for *Nezumia aequalis* and 10.03 μg/ml for *N. micronychodon* if produced by protocol (A), 1.18 μg/ml for *Nezumia aequalis* and 1.13 μg/ml for *N. micronychodon* if produced by protocol (B), and 5.35 μg/ml for *Nezumia aequalis* and 9.53 μg/ml for *N. micronychodon* if produced by protocol (C) – see Table 6 (p 138). However, during these experiments, it was recognized that it was much more important to obtain arDNA extracts

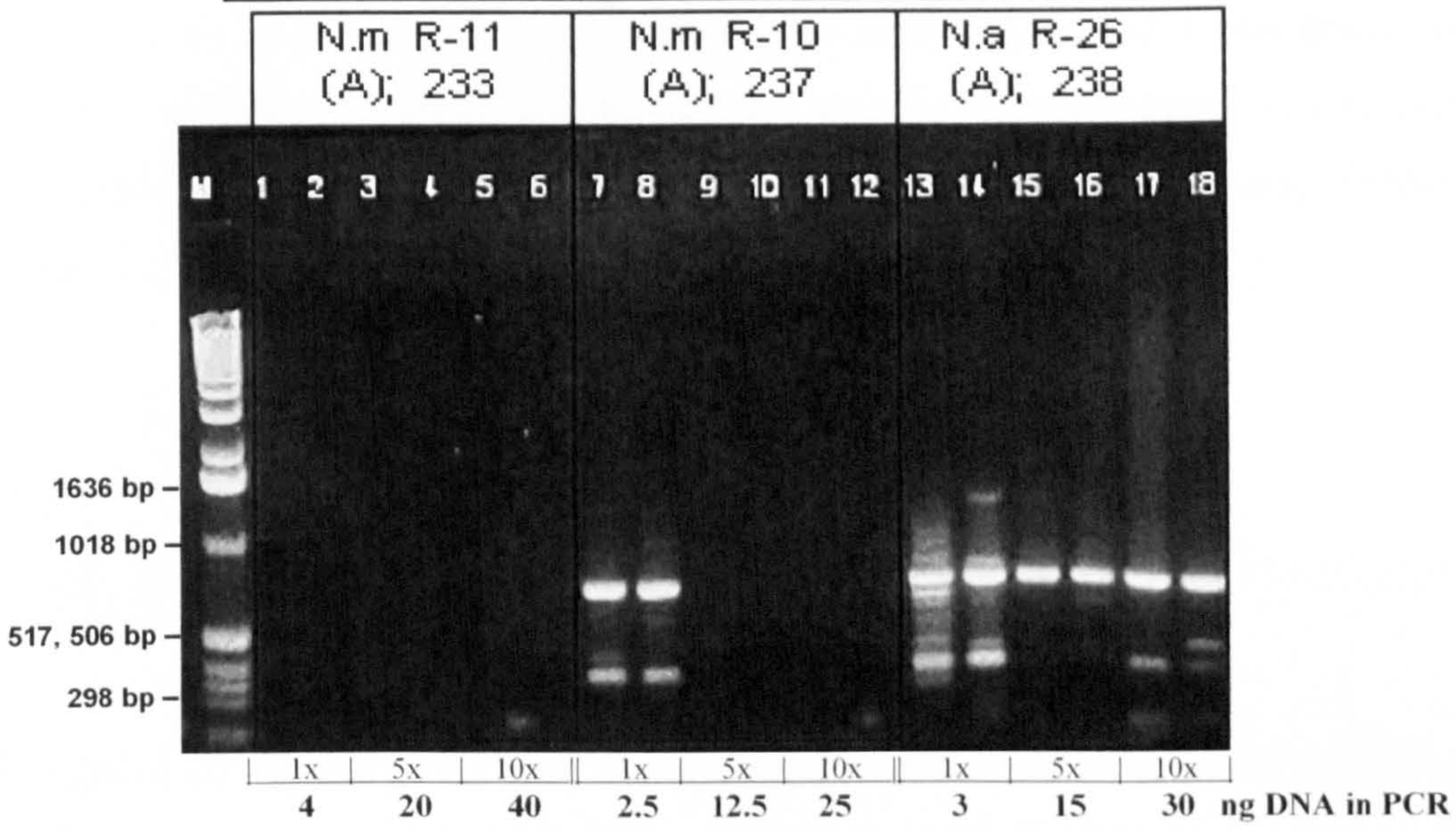
free of PCR inhibitors than trying to achieve very high yields of DNA. In other words, the production of arDNA extracts with a higher yield is less important than the production of arDNA extracts that contain a smaller amount of diffusible PCR inhibitors. The PCR inhibition caused by "diffusible" PCR inhibitors was not posing the major problem for most of arDNA extracts produced by this new protocol (A) which was not the case with previously tested protocols (especially with protocol (C)).

#### *3.1.1.1.3.1. The effects of some modifications in pre-extraction treatments*

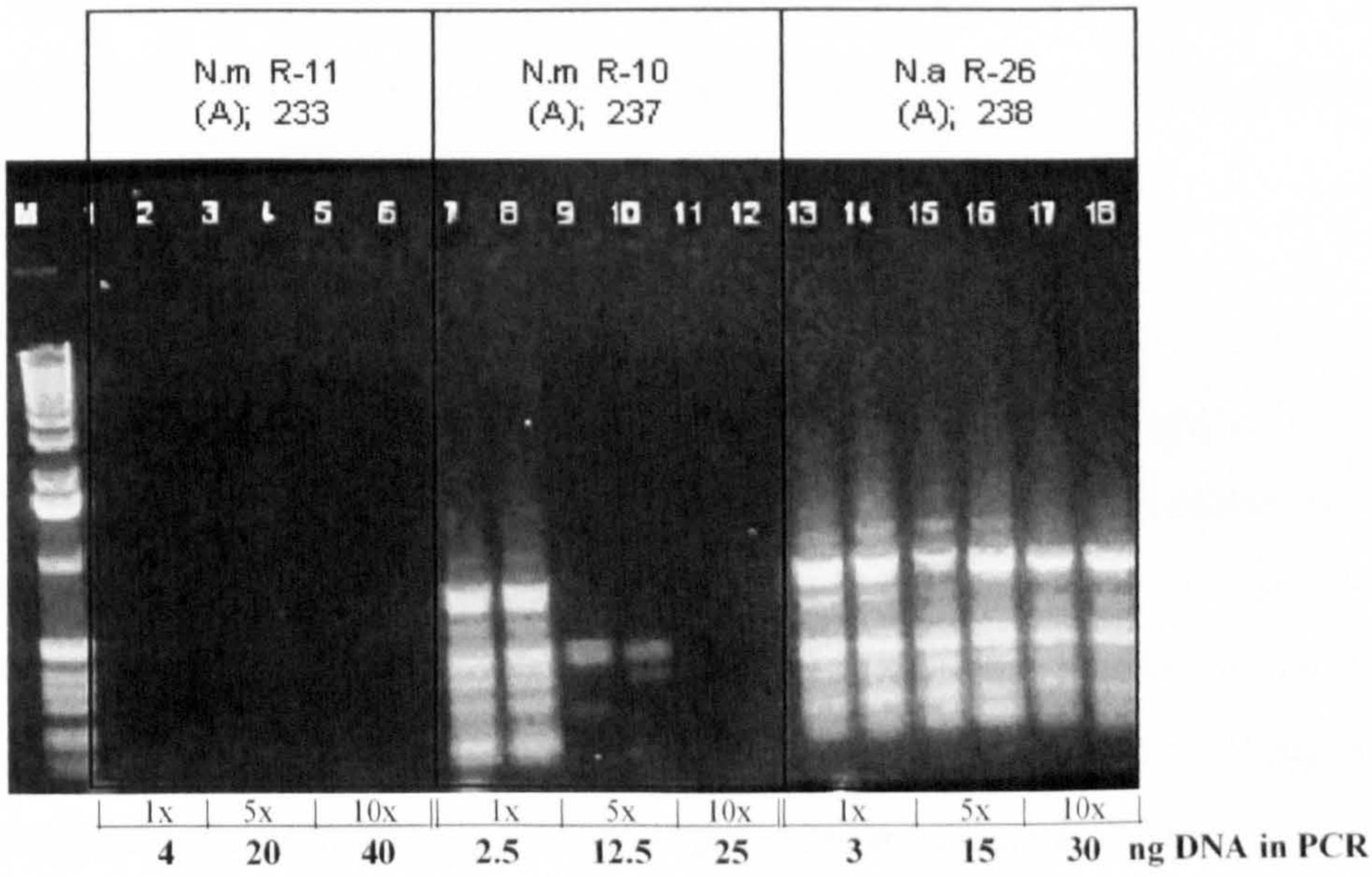
During optimisation of protocol (A), the importance was acknowledged of washing and drying regimes applied to preserved tissue. The effect of some modifications in pre-extraction treatments of protocol (A) is shown in Figure 15.

During the study, in order to speed up and simplify DNA extraction procedures, a shorter washing and drying of tissue samples was tested: washing a tissue sample in $sdH_2O$ for 10-30 min (or interchanging $sdH_2O$ and 1xGTE buffer), then 5-10 min tissue drying at room temperature. This was sufficient for some tissue samples, but better results (PCR amplifiable DNA and more reproducible PCR amplifications) were obtained with the extended pre-washing treatment (at least 24 hours) according to the experimental data obtained (see support evidence on the CD – compare, for example, Fig. S1(a)-Lanes: 15-18 and Fig. S1(h)-Lanes: 1-5 with S1(wz)-Lanes: 1-11). In a few experiments, just drying the tissue sample (at room temperature, or at 37°C) without any pre-washing, proved to be a sufficiently good pre-treatment of tissue for some arDNA extractions (but not for all tested samples and all markers; see Fig. 19(a),(b),(c)-Lanes: 6 for arDNA extract no. 206; p 130). Drying preserved tissue samples above 50°C (especially without applying tissue pre-washing) did not yield PCR amplifiable arDNAs (Fig. 15(a),(b) – Lanes: 1-6) regardless of reasonably good DNA concentration in arDNA extract (8 μg/ml in arDNA extract no. 233), or PCR performance was poor (see, for example, Fig. S1(x)-Lanes: 1-12 in support material on the CD). During the study, the full tissue pre-extraction treatment (pre-washing treatment with $sdH_2O$ and the 1xGTE buffer for 48-72 hours and drying tissue for 20-30 min at room temperature) was mostly applied on the samples studied in order to obtain formalin-arDNA extracts of good yield and quality. Also, it seems that storage of pre-washed tissue samples at –64°C, or –20°C (for few weeks/months) prior to DNA extraction procedure is beneficial in obtaining better arDNA yields and arDNA extracts more suitable for PCR.

| | | |
|---|---|---|
| **1.** Tissue sample dried at 50-60°C for 50 minutes<br>**2.** 1xGTE washing for 48 hours at room temperature<br>**3.** Tissue stored at minus 64 °C for three months<br>**4.** Tissue briefly dried at room temp. just before DNA extraction | **1.** 1xGTE washing for 24 hours at room temperature<br>**2.** Tissue dried at room temperature for four hours before DNA extraction | **1.** Interchangeably washing with sd$H_2O$ and 1xGTE for 72 hours at room temperature<br>**2.** Tissue dried at room temperature for 7.5 hours before DNA extraction |




(a) – RAPD primer 46




(b) - RAPD primer 42

**Figure 15** – The effect of tissue types and drying/washing regimes of tissue samples on the PCR performance. Tests performed with RAPD primer 46 – gel (a) and RAPD primer 42 – gel (b).

Drying tissue samples prior to DNA extraction (with and without previously applied pre-washing treatment) at different temperatures (room temperature, 37°C, ~ 50°C, or ~ 60°C) seems to be critical in obtaining amplifiable arDNA. Tissue exposure to higher temperatures (50-65°C), especially if applied without tissue pre-washing and for a prolonged period (more than 30 min), seems to have an adverse effect on DNA recovery from preserved tissue. This might be related to the situation described by Lindahl (1993b) that if the DNA is completely "dry", it will lose the double helical configuration, making the bases more vulnerable to damage from loss of structural water. The results from this study suggest that tissue drying at room temperature is probably the safest and most effective way of drying tissue as a part of pre-extraction treatment. Tissue drying by freezing is probably worth exploring further (beside the usage of dry ice as a part of "grinding technique").

#### *3.1.1.1.3.2. Application of protocol (A) on DMSO-, ethanol preserved and fresh/frozen specimens of fish*

Protocol (A) applied on **DMSO preserved tissue sample** of *Nezumia* cf. *aequalis* produced DNA extraction with a good yield (Appendix 4 – Spots V-5 and V-6; DNA concentrations 5 µg/ml and 10 µg/ml). These extracts gave good RAPD-PCR amplifications with reproducible band patterns (Fig. 16, gel (a)), good mt-PCRs (Fig. 16, gel (b)-Lanes: 1, 2), and good PCRs with RAPD-derived specific primers for *Nezumia* developed in this project (Fig. 16, gel (b)-Lane: 8).

Protocol (A) applied on control, **95% ethanol preserved tissue** samples of *Coryphaenoides armatus*, gave DNA extracts of a very good yield (Appendix 6; DNA concentrations ~20 µg/ml) and suitable for PCR applications (RAPD, and mitochondrial PCRs in particular – Fig. 17). Protocol (A) applied on control **fresh/frozen** tissue samples of cod and rainbow trout produced DNA extraction of a good yield, especially if washing regimes developed for formalin-preserved tissue samples had not been applied (DNA concentrations over 20 µg/ml; Appendix 5). PCR performances were also good (see support material on the CD).

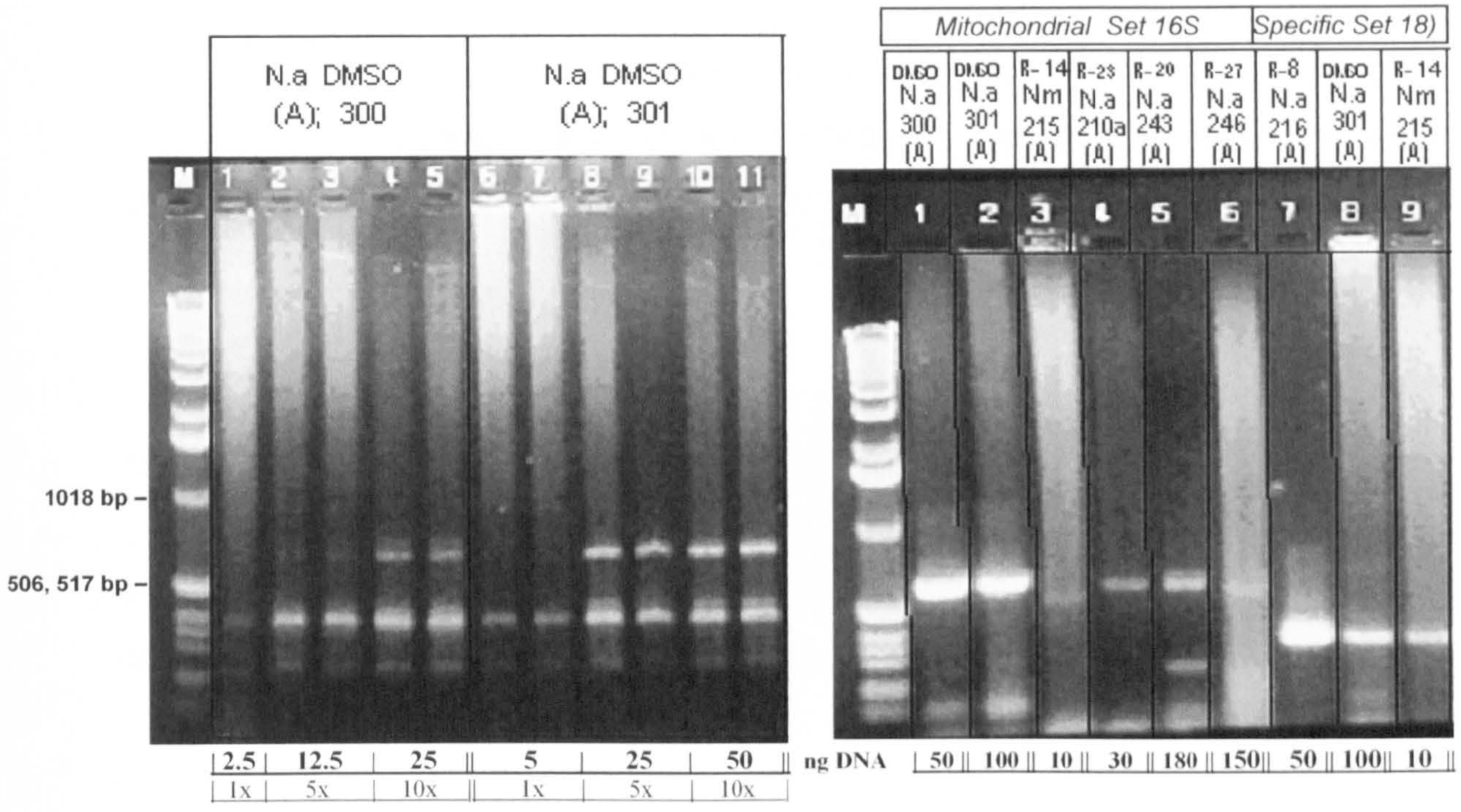



**Fig. 16 - DMSO-arDNA extracts** of *Nezumia* cf. *aequalis* (produced by protocol (A)) - tested with RAPD PCRs (gel (a)), 16S mitochondrial primer set (gel (b)-Lanes: 1, 2), and *Nezumia*-specific (RAPD-derived) set 18 (gel (b)-Lane: 8). Gel (b)-Lanes: 3-7, and 9 are PCR results with formalin-arDNA extracts produced by protocol (A).

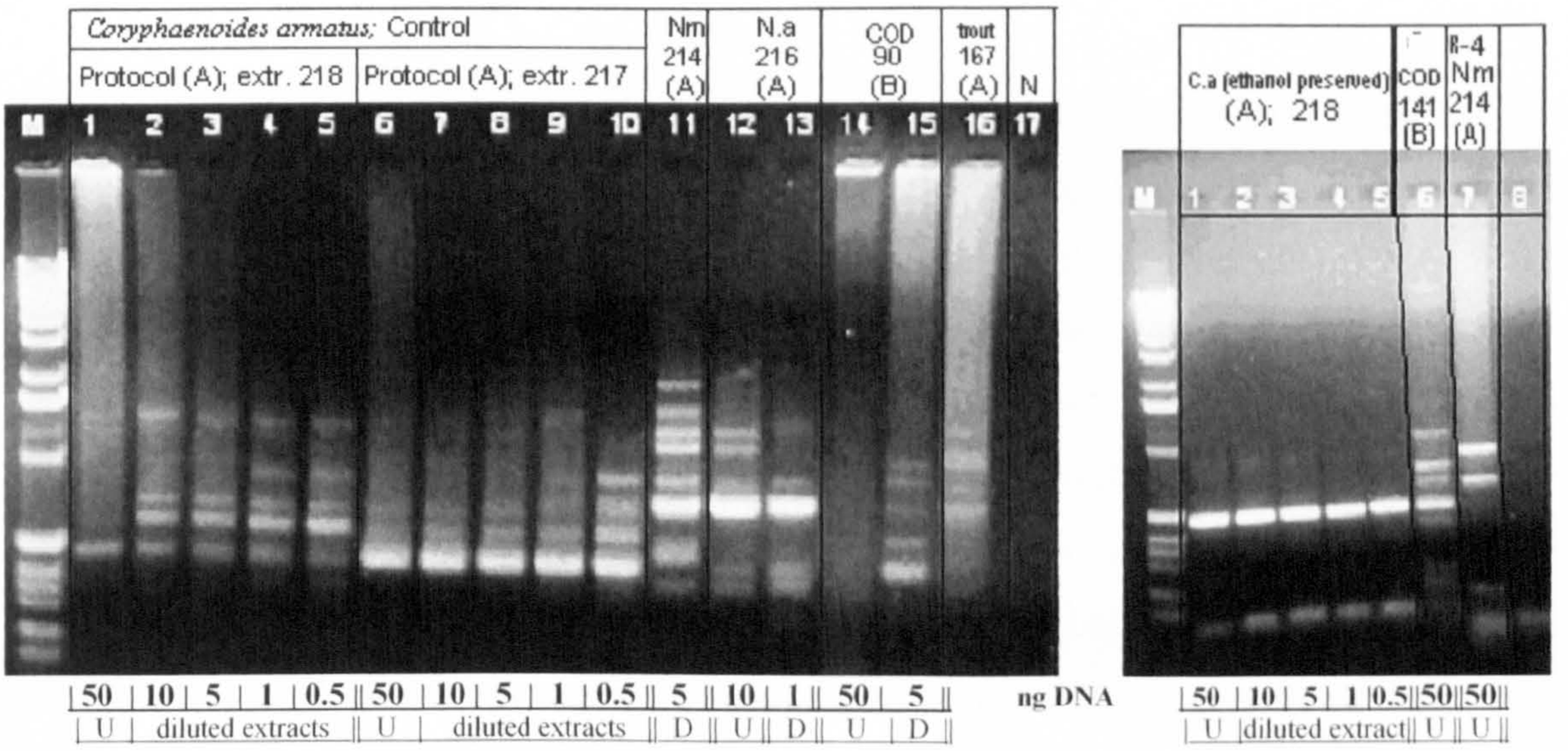



**Figure 17** – RAPD and mitochondrial PCRs with DNA extracts of 95% ethanol preserved (4 months) *Coryphaeonides armatus* (C.a), fresh/frozen cod (protocol (B)) and rainbow trout (protocol (A)), and formalin-fixed, Steedman's preserved *Nezumia aequalis* (N.a) and *N. micronychodon* (N.m). Undiluted (U) and diluted (D) DNA extracts were tested.

The data clearly demonstrate that protocol (A) is a highly suitable for extracting DNA from fish specimens, differently preserved and fresh/frozen samples. Such DNA extracts are applicable to different PCR marker systems and primers (see Tables 3-5; pp 131-134).

### *3.1.1.2. Other protocols tested/applied on preserved tissue*

A few more DNA extraction protocols were tested/used on a limited number of formalin-fixed, Steedman's preserved and ethanol preserved tissue samples of *Nezumia*.

#### 3.1.1.2.1. Protocols tested on formalin-fixed, Steedman's preserved tissue specimens of *Nezumia*

One of tested protocols was **the Promega Wizard Genomic Kit (protocol (F))** which, with small modifications in washing and grinding the tissue sample, produced good arDNA extract (DNA concentration: ~ 8 μg/ml). The arDNA extract no. 211 provided good RAPD-PCR amplifications even if very diluted arDNA samples were applied (Fig. 18 – gel (b)).

The extract 211 produced by protocol (F) also generated a faint band (mt-PCR-DNA fragment) with the 16S mitochondrial primers (Fig. 19, gel (c)-Lane: 14; p 130), indicating the possibility that further modification of this protocol might produce arDNA extracts suitable for both RAPD and mitochondrial PCR amplifications. The results of the experiment shown in Fig. 18 strongly suggest that if extracted arDNA is of a good quality and if diffusible PCR inhibitors are not greatly present in arDNA extracts, it is possible to generate reproducible RAPDs (with a particular primer) even with very low concentrations of DNA in PCRs (such as 0.4 ng of DNA in PCR – for protocol (F); Fig. 18, gel (b)-Lanes: 1, 2, 9, 10, 17, 18; or with 0.5 ng - for protocol (A) – Fig. 18, gel (a)-Lanes: 7, 8). It also suggests that a wide range of PCR-DNA template concentrations (absolute conc.: from 0.08 to 40 ng DNA, and relative conc.: from 1x to 10x) are applicable for successful RAPD-PCR amplification if arDNA extracts are of a good quality.

## RAPD primer 44

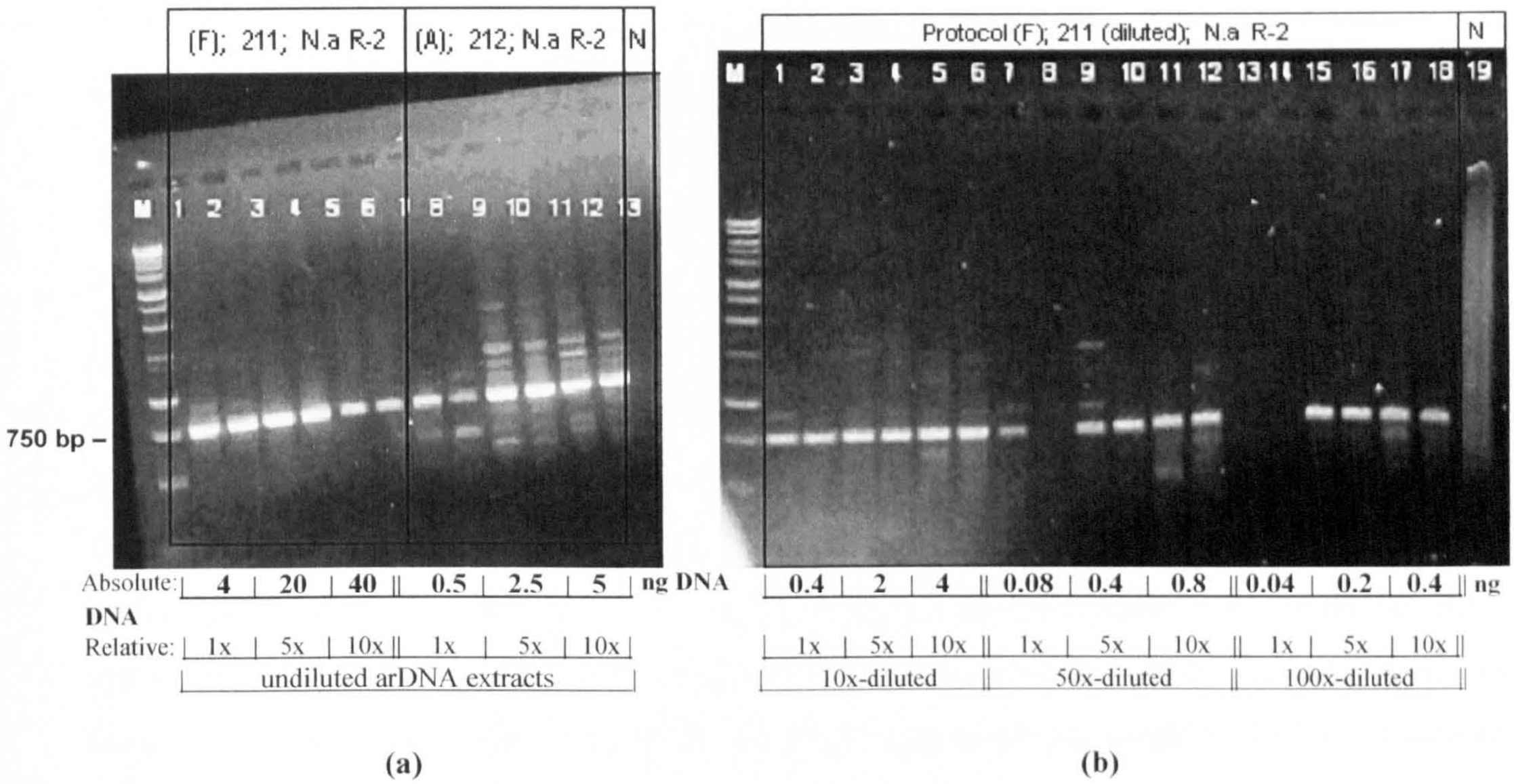



**Fig. 18** – Comparison of RAPD-PCR amplifications (primer 44) with arDNA extracts from formalin-fixed, Steedman's-preserved specimens of *Nezumia aequalis* (fish individual no. 2) produced by two different DNA extraction protocols: protocol (F) and protocol (A).

**Gel (a)** - Tested concentrations of undiluted arDNA samples in duplicates (in 25 µl PCR reaction volume) for:

- arDNA extract no. 211 produced by protocol (F) – Lanes: 1-6
- arDNA extract no. 212 produced by protocol (A) – Lanes: 7-12

**Gel (b)** - Tested concentrations of diluted arDNA extracts no. 211 produced by protocol (F); 10x-, 50x- and 100x-diluted arDNA; PCR performed in duplicates with each of tested DNA-PCR template concentrations (relative PCR-DNA template concentrations: 1x [0.5 µl of arDNA suspension (diluted or undiluted arDNA extract) in 25 µl of PCR reaction], 5x [2.5 µl of arDNA suspension (diluted or undiluted arDNA extract) in 25 µl of PCR reaction], 10x [[5 µl of arDNA suspension (diluted or undiluted arDNA extract) in 25 µl PCR reaction].

Generated RAPD band of ~750 bp was reproducible with both arDNA extracts, including diluted sample of arDNA extraction no. 211. The band was absent in PCRs with 0.04 ng of DNA (Lanes: 13 and 14) and in one with 0.08 ng of DNA in PCR reaction – Lane: 8; this is probably due to too low DNA concentration in the PCR reactions, i.e. insufficient template-amount for successful PCRs.

**"M"** indicates size marker fragments - 1Kb DNA Ladder (Promega)

**"N"** indicates negative PCR reaction

**N.a** – *Nezumia aequalis*

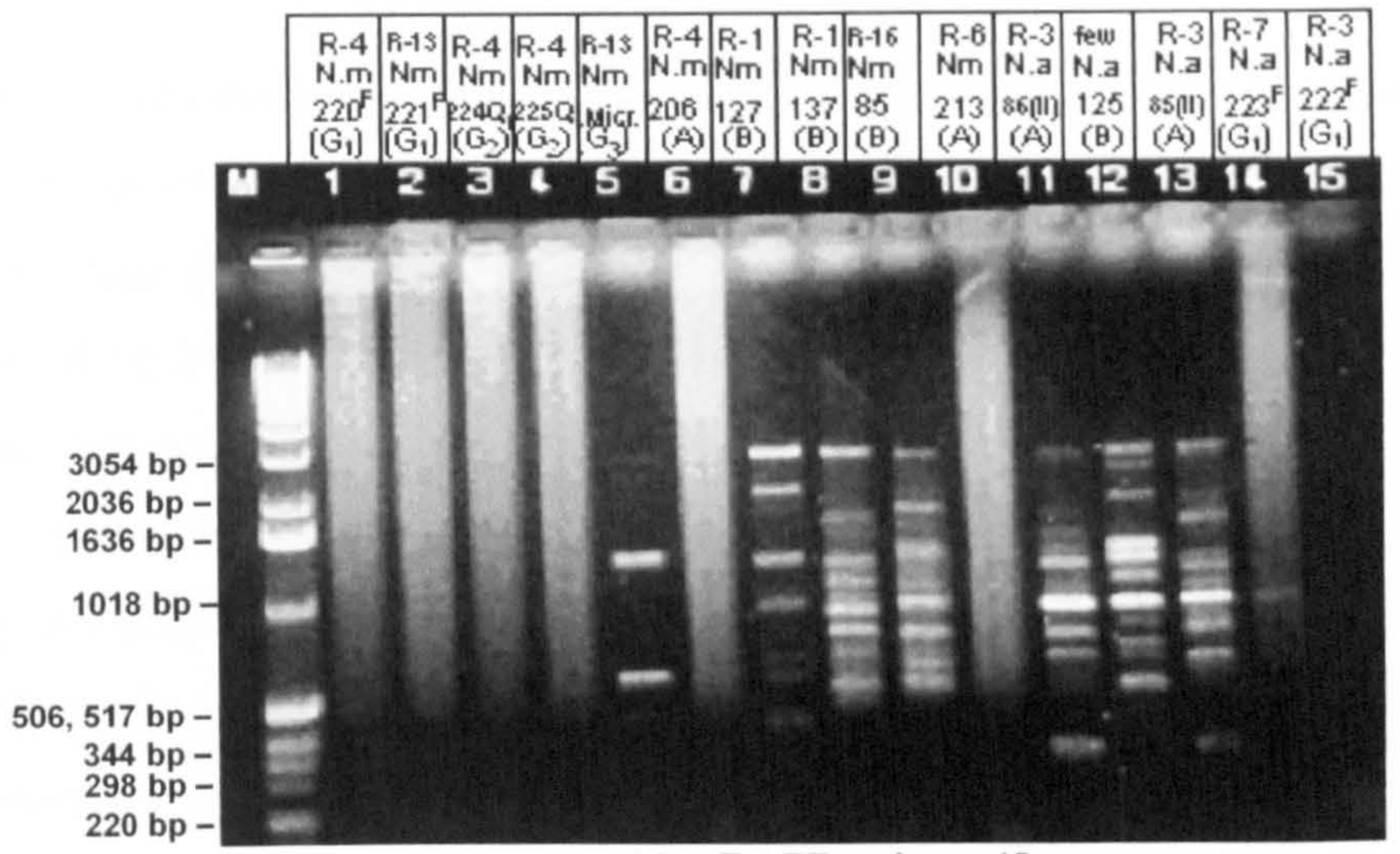
**R** - the identifier of the fish individuals (R-2 is fish individual no. 2)

**(A)** – arDNA extracts produced by protocol (A)

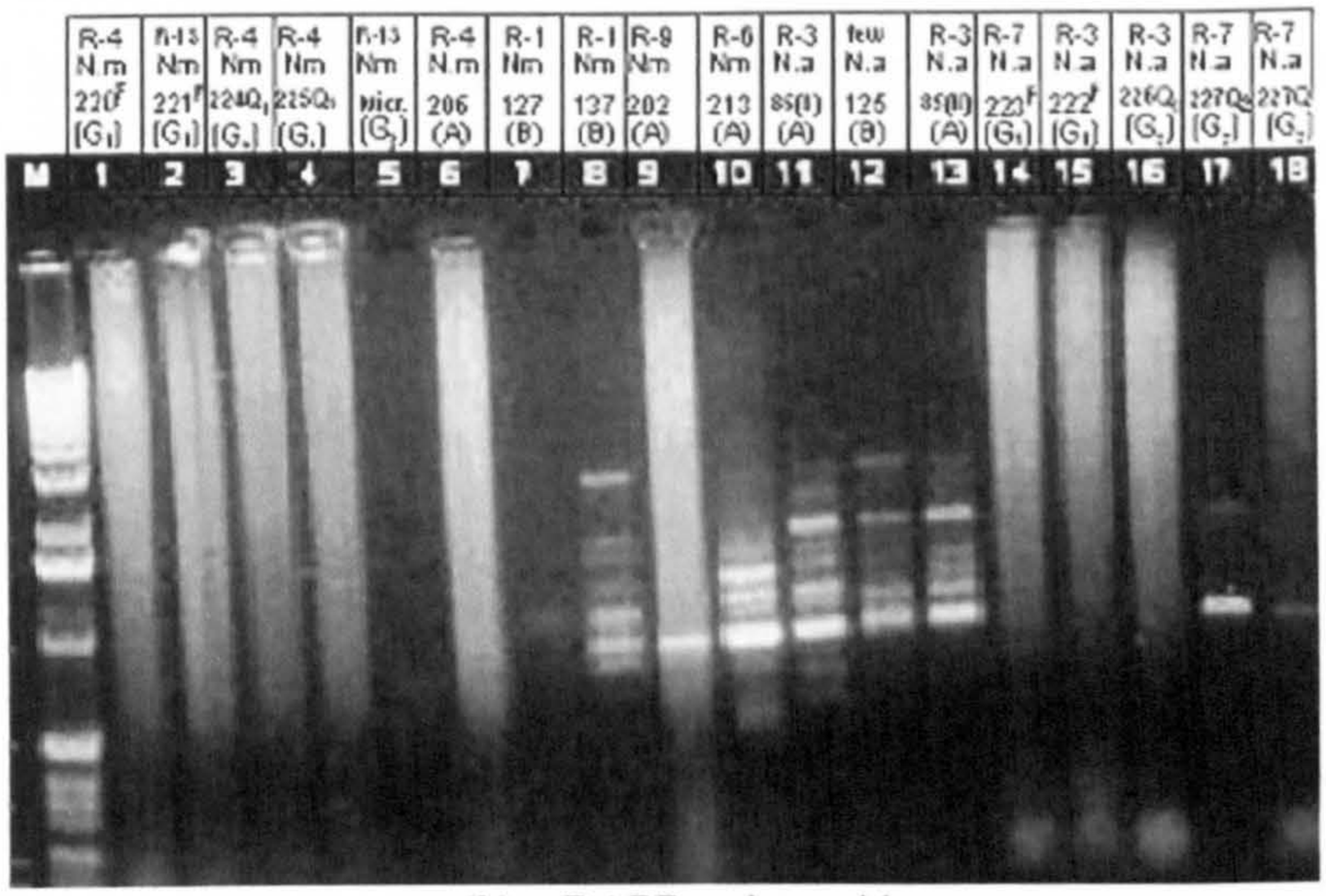
**(F)** – arDNA extracts produced by protocol (F)

**Protocol ($G_2$). The Qiagen DNeasy Tissue Kit** applied with the Southampton's very modified protocol (in comparison to the original manufacturer's protocol) was tested on four formalin-preserved tissue samples of *Nezumia aequalis* and *N. micronychodon* (both DNA elutions were tested). The yield and PCR performance with these extracts were poor, i.e. mostly without successful PCR amplification (Table 3; p 131). The results were similar with both, mitochondrial PCRs (Fig. 19, gel (c)-Lanes: 3, 4, 16, 17, 18) and RAPD-PCRs (Fig. 19, gels (a) and (b)-Lanes: 3, 4, and gel (b)-Lanes: 16, 17, 18). A faint band of appropriate size (~570 bp) was generated with 16S mitochondrial primers with one of the arDNA extracts (no. 226Q – the first elution) – Fig 19, gel (c)-Lane: 16, but not with RAPD (gel (b)-Lane: 16). Another arDNA extract (no. 227Q – both elutions) generated RAPDs - Fig 19, gel (b)-Lanes: 17 and 18, but not mitochondrial PCR product (gel (c)-Lanes: 17 and 18). The applied protocol ($G_2$) was adapted for investigations of short-term ethanol preserved organisms by the Southampton research team, and therefore it is worthwhile to test this Kit again applying the original protocol and by making modifications that would be more appropriate for formalin-preserved specimens. The data from this study indicated that this Qiagen commercial Kit with appropriate modifications might be useful in DNA extractions from formalin-fixed fish specimens.

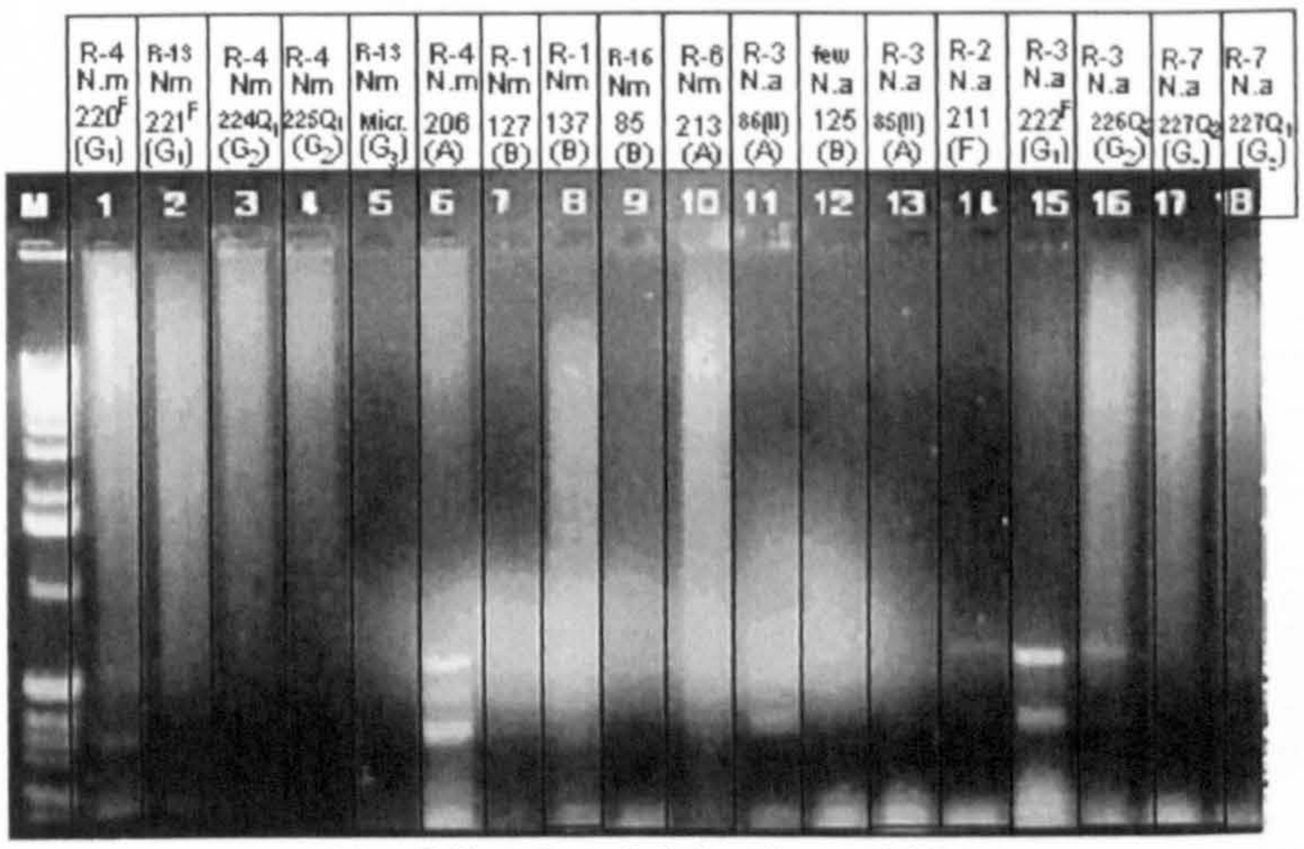
**The MicroLYSIS commercial kit, protocol ($G_3$),** produced a successful RAPD-PCR amplification with RAPD primer 48, but not with primer 44 (Fig. 19, gels (a) and (b), compare lanes 5), and not with mitochondrial 16S primers (gel (c)-Lane 5). Because of the protocol simplicity, it was not expected that it would be possible to extract PCR amplifiable DNA from formalin-fixed specimens. It was assumed that the treatments undertaken had insufficient "strength" for DNA extraction from preserved tissue. The PCR results indicated the value of further investigation of this protocol (speed, simplicity and cost-effectiveness). Using this klt as it is, it is not applicable for effective DNA extractions from preserved specimens, but combining it with some more "powerful" DNA extraction steps and/or DNA extraction protocols, it will be worthwhile testing.




**(a) – RAPD primer 48**




**(b) – RAPD primer 44**




**(c) – Mitochondrial primer 16S**

**Fig. 19** – Archival DNA extracts produced by different DNA extraction protocols - tested with two RAPD primers and with mitochondrial set for the 16S gene.

**Protocol ($G_1$).** Besides protocol (C), one more phenol-based protocol (protocol ($G_1$)) was tested on formalin-fixed specimens of *Nezumia*. Protocol ($G_1$) exhibited better PCR performance than protocol (C), but still most RAPD-PCRs did not yield PCR products (Table 3), or RAPD bands were of a low intensity with strong background smearing (Fig. 19, gels (a) and (b): compare lanes: 1, 2, 14, 15). Some arDNA extracts produced by protocol ($G_1$) gave successful mitochondrial PCR amplifications (Fig. 19, gel (c)-Lane: 15). It is interesting to observe that some arDNA extracts produced by protocol ($G_1$) did not yield RAPD-PCR products (or if they did, RAPD profiling was of a poor quality with bands of a low intensity), but they yielded PCR product(s) from the mitochondrial genome. An example of this is arDNA extract no. $222^F$ which gave unsuccessful (or very poor quality) RAPD-PCRs whereas mitochondrial PCR of the 16S gene region was reasonable successful (Fig. 19, compare lanes 15 on gels (a), (b), and (c)). Extractions $220^F$ and $221^F$ were unsuccessful with both PCR marker systems (Fig. 19, gels (a), (b), and (c): compare lanes 1 and 2), although extraction no. $220^F$ gave a successful mitochondrial PCR amplification to some extent (Fig. 19, gel (c)-Lane: 1). Overall, this protocol produced formalin-Steedman-arDNA with poor and inconsistent RAPD-PCR performances, but their application to the mitochondrial PCR marker system was successful to some extent (Table 3).

**Table 3** – *PCR success rate (selected mitochondrial and RAPD PCRs) related to DNA extraction method applied to formalin-fixed, Steedman's preserved specimens of Nezumia*

## a) mitochondrial PCRs*

| | PCR target | Protocols (A) | (B) | (C) | (D) | (F) | (G1) | (G2) | (G3) |
|---|---|---|---|---|---|---|---|---|---|
| | **mt COIII (550 bp)** | | | | | | | | |
| **N.m** | PCRs (successful/ total) (% successful) | 3/ 7 (43%) | 1/ 16 (6%) | 0/ 35 (0%) | 0/ 3 (0%) | n/a | n/a | n/a | n/a |
| **N.a** | PCRs (successful/ total) (% successful) | 1/ 4 (25%) | 0/ 9 (0%) | 0/ 30 (0%) | 0/ 3 (0%) | 0/ 2 (0%) | n/a | n/a | n/a |
| **N.m + N.a** | **PCRs (successful/ total) (% successful)** | **4/ 11 (36%)** | **1/ 25 (4%)** | **0/ 65 (0%)** | **0/ 6 (46%)** | **0/ 2 (0%)** | n/a | n/a | n/a |
| | **mt 16S (570 bp)** | | | | | | | | |
| **N.m** | PCRs (successful/ total) (% successful) | 5/ 27 (18.5%) | 2/ 7 (28.5%) | n/a | 0/4 (0%) | n/a | 2/ 4 (50%) | 0/ 6 (0%) | 0/ 2 (0%) |
| **N.a** | PCRs (successful/ total) (% successful) | 9/ 38 (24%) | 0/ 1 (0%) | n/a | n/a | 1/ 5 (20%) | 7/ 8 (87.5%) | 1/ 8 (12.5%) | n/a |
| **N.m + N.a** | **PCRs (successful/ total) (% successful)** | **14/ 65 (21.5%)** | **2/ 8 (25%)** | **n/a** | **0/ 4 (0%)** | **1/ 5 (20%)** | **9/ 12 (75%)** | **1/ 14 (7%)** | **0/ 2 (0%)** |
| | **2 mt primer sets (COIII + 16S)** | | | | | | | | |
| **N.m** | PCRs (successful/ total) (% successful) | 8/ 34 (23.5%) | 3/ 23 (13%) | 0/ 35 (0%) | 0/ 7 (0%) | n/a | 2/ 4 (50%) | 0/ 6 (0%) | 0/ 2 (0%) |
| **N.a** | PCRs (successful/ total) (% successful) | 10/ 42 (24%) | 0/ 10 (0%) | 0/ 30 (0%) | 0/ 3 (0%) | 1/ 7 (14%) | 7/ 8 (87.5%) | 1/ 8 (12.5%) | n/a |
| **N.m + N.a** | **PCRs (successful/ total) (% successful)** | **18/ 76 (24%)** | **3/ 33 (9%)** | **0/ 65 (0%)** | **0/ 10 (0%)** | **1/ 7 (14%)** | **9/ 12 (75%)** | **1/ 14 (7%)** | **0/ 2 (0%)** |

# b) RAPD PCRs*

| | | Protocols (A) | (B) | (C) | (D) | (F) | (G1) | (G2) | (G3) |
|---|---|---|---|---|---|---|---|---|---|
| | **RAPD Primer 11** | | | | | | | | |
| **N.m** | PCRs (successful/ total) (% successful) | 29/ 84 (34.5%) | 1/ 2 (50%) | n/a | n/a | n/a | n/a | n/a | n/a |
| | No. of bands (means) | 3.58 | 1.00 | n/a | n/a | n/a | n/a | n/a | n/a |
| **N.a** | PCRs (successful/ total) (% successful) | 69/ 108 (64%) | 0/ 2 (0%) | n/a | n/a | n/a | n/a | n/a | n/a |
| | No. of bands (means) | 5.00 | 0.00 | n/a | n/a | n/a | n/a | n/a | n/a |
| **N.m + N.a** | **PCRs (successful/ total) (% successful)** | **98/ 192 (51%)** | **1/ 4 (25%)** | **n/a** | **n/a** | **n/a** | **n/a** | **n/a** | **n/a** |
| | **No. of bands (means)** | **4.58** | **1.00** | **n/a** | **n/a** | **n/a** | **n/a** | **n/a** | **n/a** |
| | **RAPD Primer 15** | | | | | | | | |
| **N.m** | PCRs (successful/ total) (% successful) | 55/ 89 (62%) | 9/ 9 (100%) | n/a | n/a | n/a | n/a | n/a | n/a |
| | No. of bands (means) | 3.68 | 5.67 | n/a | n/a | n/a | n/a | n/a | n/a |
| **N.a** | PCRs (successful/ total) (% successful) | 95/ 108 (88%) | 8/ 8 (100%) | n/a | n/a | n/a | n/a | n/a | n/a |
| | No. of bands (means) | 4.98 | 3.88 | n/a | n/a | n/a | n/a | n/a | n/a |
| **N.m + N.a** | **PCRs (successful/ total) (% successful)** | **150/ 197 (76%)** | **17/ 17 (100%)** | **n/a** | **n/a** | **n/a** | **n/a** | **n/a** | **n/a** |
| | **No. of bands (means)** | **4.50** | **4.82** | **n/a** | **n/a** | **n/a** | **n/a** | **n/a** | **n/a** |
| | **RAPD Primer 29** | | | | | | | | |
| **N.m** | PCRs (successful/ total) (% successful) | 31/ 116 (27%) | 69/ 154 (45%) | 0/ 12 (0%) | 2/ 11 (18%) | n/a | n/a | n/a | n/a |
| | No. of bands (means) | 4.25 | 3.12 | 0.00 | 1.00 | n/a | n/a | n/a | n/a |
| **N.a** | PCRs (successful/ total) (% successful) | 26/ 92 (28%) | 8/ 78 (10%) | 0/ 11 (0%) | n/a | n/a | n/a | n/a | n/a |
| | No. of bands (means) | 5.54 | 1.67 | 0.00 | n/a | n/a | n/a | n/a | n/a |
| **N.m + N.a** | **PCRs (successful/ total) (% successful)** | **57/ 208 (27%)** | **77/ 232 (33%)** | **0/ 23 (0%)** | **2/ 11 (18%)** | **n/a** | **n/a** | **n/a** | **n/a** |
| | **No. of bands (means)** | **4.83** | **2.96** | **0.00** | **1.00** | **n/a** | **n/a** | **n/a** | **n/a** |
| | **RAPD Primer 44** | | | | | | | | |
| **N.m** | PCRs (successful/ total) (% successful) | 53/ 117 (45%) | 163/ 265 (62%) | 13/ 59 (22%) | 15/ 20 (75%) | n/a | 0/ 2 (0%) | 0/ 2 (0%) | 1/ 1 (100%) |
| | No. of bands (means) | 3.98 | 1.79 | 3.20 | 2.20 | n/a | 0.00 | 0.00 | 2.00 |
| **N.a** | PCRs (successful/ total) (% successful) | 165/ 218 (76%) | 151/ 278 (54%) | 22/ 61 (36%) | 12/ 24 (50%) | 21/ 24 (87.5%) | 0/ 2 (0%) | 2/ 4 (50%) | n/a |
| | No. of bands (means) | 4.90 | 4.14 | 3.09 | 1.10 | 1.90 | 0.00 | 3.50 | n/a |
| **N.m + N.a** | **PCRs (successful/ total) (% successful)** | **218/ 335 (65%)** | **314/ 543 (58%)** | **35/ 120 (29%)** | **27/ 44 (61%)** | **21/ 24 (87.5%)** | **0/ 4 (0%)** | **2/ 6 (33%)** | **1/ 1 (100%)** |
| | **No. of bands (means)** | **4.68** | **3.87** | **3.14** | **1.70** | **1.90** | **0.00** | **3.50** | **2.00** |
| | **RAPD Primer 46** | | | | | | | | |
| **N.m** | PCRs (successful/ total) (% successful) | 69/ 132 (52%) | 3/ 3 (100%) | n/a | n/a | n/a | n/a | n/a | n/a |
| | No. of bands (means) | 3.01 | 1.00 | n/a | n/a | n/a | n/a | n/a | n/a |
| **N.a** | PCRs (successful/ total) (% successful) | 133/ 162 (82%) | 0/ 4 (0%) | n/a | n/a | n/a | n/a | n/a | n/a |
| | No. of bands (means) | 4.07 | 0.00 | n/a | n/a | n/a | n/a | n/a | n/a |
| **N.m + N.a** | **PCRs (successful/ total) (% successful)** | **202/ 294 (69%)** | **3/ 7 (43%)** | **n/a** | **n/a** | **n/a** | **n/a** | **n/a** | **n/a** |
| | **No. of bands (means)** | **3.70** | **1.00** | **n/a** | **n/a** | **n/a** | **n/a** | **n/a** | **n/a** |
| | **RAPD Primer 48** | | | | | | | | |
| **N.m** | PCRs (successful/ total) (% successful) | 31/ 85 (36.5%) | 11/ 11 (100%) | n/a | n/a | n/a | 0/ 2 (0%) | 0/ 2 (0%) | 1/ 1 (100%) |
| | No. of bands (means) | 3.22 | 4.32 | n/a | n/a | n/a | 0.00 | 0.00 | 4.00 |
| **N.a** | PCRs (successful/ total) (% successful) | 62/ 116 (53%) | 10/ 10 (100%) | n/a | n/a | 4/ 6 (67%) | 1/ 2 (50%) | 1/ 4 (25%) | n/a |
| | No. of bands (means) | 5.86 | 4.25 | n/a | n/a | 7.38 | 2.00 | 10.00 | n/a |
| **N.m + N.a** | **PCRs (successful/ total) (% successful)** | **93/ 201 (46%)** | **21/ 21 (100%)** | **n/a** | **n/a** | **4/ 6 (67%)** | **1/ 4 (25%)** | **1/ 6 (17%)** | **1/ 1 (100%)** |
| | **No. of bands (means)** | **4.80** | **4.29** | **n/a** | **n/a** | **7.38** | **2.00** | **10.00** | **4.00** |
| | **6 RAPD primers (11+15+29+44+46+48)** | | | | | | | | |
| **N.m** | PCRs (successful/ total) (% successful) | 268/ 623 (43%) | 256/ 444 (57.5%) | 13/ 71 (18%) | 17/ 31 (55%) | n/a | 0/ 4 (0%) | 0/ 4 (0%) | 2/ 2 (100%) |
| | No. of bands (means) | 3.56 | 3.54 | 3.23 | 2.03 | n/a | 0.00 | 0.00 | 3.00 |
| **N.a** | PCRs (successful/ total) (% successful) | 550/ 804 (68%) | 177/ 380 (46.5%) | 22/ 72 (30.5%) | 12/ 24 (50%) | 25/ 30 (83%) | 1/ 4 (25%) | 3/ 8 (37.5%) | n/a |
| | No. of bands (means) | 4.87 | 4.00 | 3.09 | 1.08 | 2.78 | 2.00 | 5.67 | n/a |
| **N.m + N.a** | **PCRs (successful/ total) (% successful)** | **818/ 1427 (57%)** | **433/ 824 (52.5%)** | **35/ 143 (24.5%)** | **29/ 55 (53%)** | **25/ 30 (83%)** | **1/ 8 (12.5%)** | **3/ 12 (25%)** | **2/ 2 (100%)** |
| | **No. of bands (means)** | **4.42** | **3.73** | **3.14** | **1.64** | **2.78** | **2.00** | **5.67** | **3.00** |

**N.a** – *Nezumia aequalis*
**N.m** – *Nezumia micronychodon*
**Nm + N.a** – Both species observed together
**n/a** – not aplicable (no data with that DNA extraction protocol)

* see Table 4 for statistical significance related to the PCR success rate and DNA extracted with a particular protocol;

see Table 5 for statistical significance related to the number of RAPD bands generated with arDNA extracts produced by a specific DNA extraction protocol

**Table 4** – *Statistical significance related to the PCR success rate related to the extraction protocol applied* (see Table 3)

| | | $\chi^2$ | p | estimate | $\chi^2$ | p | estimate | $\chi^2$ | p | estimate | $\chi^2$ | p | estimate |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Protocol | | | (B) | | | (C) | | | (D) | | | (F) | |
| | (A) | 3.330 | 0.068 | -0.268 | 42.435 | < 0.001 | -1.508 | 0.090 | 0.770 | -0.100 | 1.530 | 0.216 | 0.732 |
| *RAPD* | (B) | | | | 35.172 | < 0.001 | -1.244 | 0.138 | 0.710 | 0.120 | 10.560 | 0.001 | 1.716 |
| *Primer 44* | (C) | | | | | | | 13.808 | < 0.001 | 1.346 | 19.980 | < 0.001 | 2.518 |
| | (D) | | | | | | | | | | 8.268 | 0.004 | 1.946 |
| | (A) | 1.720 | 0.190 | -0.118 | 57.950 | < 0.001 | -1.464 | 0.032 | 0.858 | -0.051 | 3.200 | 0.074 | 0.817 |
| *6 RAPD* | (B) | | | | 42.452 | < 0.001 | -1.266 | 0.000 | 0.989 | -0.004 | 15.210 | < 0.001 | 1.696 |
| *Primers* | (C) | | | | | | | 15.187 | < 0.001 | 1.305 | 25.110 | < 0.001 | 2.430 |
| | (D) | | | | | | | | | | 6.968 | 0.008 | 1.609 |
| | (A) | 3.925 | 0.048 | -1.217 | 25.110 | < 0.001 | -18.000 | 5.074 | 0.024 | -16.000 | 0.244 | 0.622 | -0.540 |
| *2 mt* | (B) | | | | 5.802 | 0.016 | -16.000 | 1.693 | 0.193 | -15.000 | 2.075 | 0.150 | 17.000 |
| *Primer sets* | (C) | | | | | | | 0.000 | 0.999 | 2.000 | 3.732 | 0.530 | 18.000 |
| | (D) | | | | | | | | | | 0.872 | 0.350 | 16.000 |

**Table 5** – *Statistical significance related to the number of RAPD bands – in conjunction with an extraction protocol* (see Table 3)

**RAPD Primer 44**

Estimates of pairwise differences (with the estimates of the means on the diagonal)

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| (A) | 4.039 | | | | | | | | |
| (B) | 0.647 | 3.391 | | | | | | | |
| (C) | 1.726 | 1.079 | 2.312 | | | | | | |
| (D) | 2.681 | 2.034 | 0.955 | 1.357 | | | | | |
| (F) | 2.198 | 1.551 | 0.472 | -0.483 | 1.840 | | | | |
| (G1) | 4.039 | 3.391 | 2.312 | 1.357 | 1.840 | 0.000 | | | |
| (G2) | 3.164 | 2.516 | 1.438 | 0 482 | 0.965 | -0.875 | 0.875 | | |
| (G3) | 1.979 | 1.332 | 0.253 | -0.702 | -0.219 | -2.060 | -1.185 | 2.060 | |
| Mix | * | * | * | * | * | * | * | * | * |
| Protocol | (A) | (B) | (C) | (D) | (F) | (G1) | (G2) | (G3) | Mix |

t values of pairwise differences (t)

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| (A) | * | | | | | | | | |
| (B) | 0.70 | * | | | | | | | |
| (C) | 1.88 | 1.17 | * | | | | | | |
| (D) | 2.92 | 2.21 | 1.04 | * | | | | | |
| (F) | 1.90 | 1.34 | 0.41 | -0.42 | * | | | | |
| (G1) | 4.39 | 3.69 | 2.51 | 1.48 | 1.59 | * | | | |
| (G2) | 3.44 | 2.74 | 1.56 | 0.52 | 0.83 | -0.95 | * | | |
| (G3) | 1.71 | 1.15 | 0.22 | -0.61 | -0.16 | -1.78 | -1.02 | * | |
| Mix | * | * | * | * | * | * | * | * | * |
| Protocol | (A) | (B) | (C) | (D) | (F) | (G1) | (G2) | (G3) | Mix |

t probabilities of pairwise differences (p)

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| (A) | * | | | | | | | | |
| (B) | 0.513 | * | | | | | | | |
| (C) | 0.119 | 0.294 | * | | | | | | |
| (D) | 0.033 | 0.078 | 0.347 | * | | | | | |
| (F) | 0.116 | 0.238 | 0.700 | 0.694 | * | | | | |
| (G1) | 0.007 | 0.014 | 0.054 | 0.200 | 0.173 | * | | | |
| (G2) | 0.018 | 0.041 | 0.179 | 0.622 | 0.442 | 0.385 | * | | |
| (G3) | 0.148 | 0.302 | 0.835 | 0.571 | 0.882 | 0.135 | 0.353 | * | |
| Mix | * | * | * | * | * | * | * | * | * |
| Protocol | (A) | (B) | (C) | (D) | (F) | (G1) | (G2) | (G3) | Mix |

**6 RAPD primers**

Estimates of pairwise differences (with the estimates of the means on the diagonal)

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| (A) | 3.487 | | | | | | | | |
| (B) | 0.537 | 2.950 | | | | | | | |
| (C) | 2.598 | 2.060 | 0.890 | | | | | | |
| (D) | 2.232 | 1.695 | -0.365 | 1.255 | | | | | |
| (F) | -0.901 | -1.439 | -3.499 | -3.134 | 4.389 | | | | |
| (G1) | 3.237 | 2.700 | 0.640 | 1.005 | 4.139 | 0.250 | | | |
| (G2) | 1.800 | 1.263 | -0.798 | -0.432 | 2.701 | -1.438 | 1.687 | | |
| (G3) | 0.236 | -0.301 | -2.362 | -1.996 | 1.137 | -3.001 | -1.564 | 3.251 | |
| Mix | * | * | * | * | * | * | * | * | * |
| Protocol | (A) | (B) | (C) | (D) | (F) | (G1) | (G2) | (G3) | Mix |

t values of pairwise differences (t)

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| (A) | * | | | | | | | | |
| (B) | 0.45 | * | | | | | | | |
| (C) | 2.17 | 1.72 | * | | | | | | |
| (D) | 1.87 | 1.42 | -0.31 | * | | | | | |
| (F) | -0.60 | -0.96 | -2.32 | -2.08 | * | | | | |
| (G1) | 2.71 | 2.26 | 0.53 | 0.84 | 2.75 | * | | | |
| (G2) | 1.50 | 1.06 | -0.67 | -0.36 | 1.79 | -1.20 | * | | |
| (G3) | 0.16 | -0.20 | -1.57 | -1.33 | 0.62 | -1.99 | -1.04 | * | |
| Mix | * | * | * | * | * | * | * | * | * |
| Protocol | (A) | (B) | (C) | (D) | (F) | (G1) | (G2) | (G3) | Mix |

t probabilities of pairwise differences (p)

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| (A) | * | | | | | | | | |
| (B) | 0.672 | * | | | | | | | |
| (C) | 0 082 | 0.146 | * | | | | | | |
| (D) | 0.121 | 0.216 | 0.772 | * | | | | | |
| (F) | 0.576 | 0.383 | 0.068 | 0.092 | * | | | | |
| (G1) | 0.043 | 0.074 | 0.616 | 0.439 | 0.040 | * | | | |
| (G2) | 0.193 | 0.340 | 0.534 | 0.732 | 0.133 | 0.283 | * | | |
| (G3) | 0.881 | 0.849 | 0.178 | 0.242 | 0.561 | 0.103 | 0.347 | * | |
| Mix | * | * | * | * | * | * | * | * | * |
| Protocol | (A) | (B) | (C) | (D) | (F) | (G1) | (G2) | (G3) | Mix |

### 3.1.1.2.2. Protocols tested on ethanol preserved tissue specimens of *Nezumia*

**Phenol-based protocol (H)** was tested only on *Nezumia* short-term ethanol preserved tissue samples (tissue stored in ethanol 1-1.5 yrs). This protocol (H) produced a high-molecular-weight DNA (Fig. 20, gels (a) and (b)) and provided a good yield of DNA (DNA concentrations were from 2 μg/ml to over 20 μg/ml). The PCRs were successful with four (out of five) ethanol-arDNA extracts (Fig. 20, gel (c)). One ethanol-arDNA extract (no. NHM-2) did not produce PCR products with any of the tested PCR marker systems and primers (e.g., Fig. 20, gel (c)-Lanes: 2, 8, 14; marked with red arrows). This demonstrates that some of the chemicals used in this phenol-based DNA extraction protocol might have adverse effects even on extracting DNA from short period ethanol-preserved tissue. It is presumed that a NHM-2 extract contained strong inhibitor(s) causing the PCR to fail. The DNA concentration of 10 ng in the PCR reaction was more than sufficient for successful PCR amplifications with such a good quality of DNA (but also not being in excess that can cause the PCR failure). Assumed PCR inhibition was so strong that diluting a DNA sample of the NHM-2 DNA extract did not aid in producing successful PCR amplifications. The additional evidence on mackerel (see sections 3.3.1. and 3.4.2.) indicated that sodium acetate, used for DNA precipitation, was the most likely chemical that caused the problem. Also, three ethanol-arDNA extracts of *Nezumia* (NHM-3, NHM-4 and NHM-5), which were kept only for 45 min (instead of overnight) in a fridge at 4°C for DNA precipitation, did not exhibit PCR inhibition (or not so strong PCR inhibition that would cause unsuccessful PCR amplifications; see support material on the CD). This indicates the importance of each step in producing PCR amplifiable arDNA.

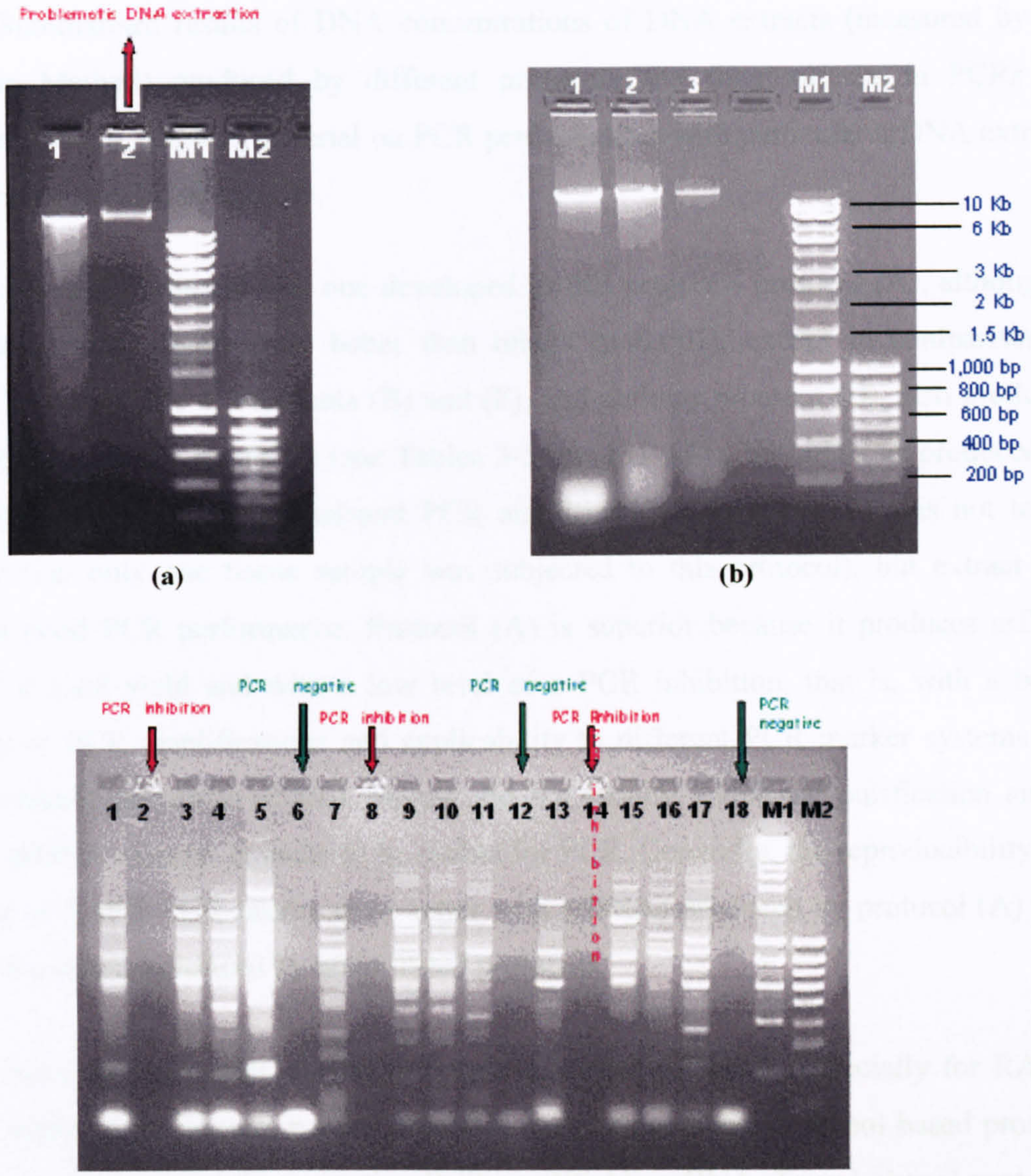



**Figure 20** – **The PCR inhibition** with one ethanol-arDNA extract produced by phenol protocol (H). The red arrows indicate the problematic DNA extract (NHM-2) of *N. aequalis* (gel (a)–Lane 2) and inhibited PCRs (gel (c)–Lanes: 2, 8, and 14). Green arrows indicate PCR negatives ($sdH_2O$ instead of DNA was used as a PCR template).
"$M_1$" indicates quantitative size marker fragments - HyperLadder I (Bioline)
"$M_2$" - Bioline HyperLadder IV (Bioline).

Generally, protocol (H) exhibited inconsistency in producing PCR amplifiable DNA even from ethanol preserved tissue samples. The data also indicate a high sensitivity to the smallest modifications to the extraction procedure (duration of DNA precipitations in these experiments) in producing amplifiable DNA. These could be problematic for usage on formalin preserved collections, because such samples require more robustness in DNA extraction protocols applied.

In this project, 324 DNA extracts were produced by using nine different DNA extraction protocols. Summarised results of DNA concentrations of DNA extracts (measured by the Saran Wrap Method) produced by different protocols and their success in PCRs are presented in Table 6. Support material on PCR performances with particular arDNA extracts and markers is provided on the CD.

The most successful protocol was one developed in this project – protocol (A), although it was not statistically significantly better than others ($p<0.001$), except in comparison to protocol (C) – see Table 4. Protocols (B) and (F), and perhaps protocol (D), also produced arDNA extracts suitable for PCRs (see Tables 3-5; pp 131-134), but arDNAs produced by protocols (B) and (D) gave inconsistent PCR amplifications. Protocol (F) was not tested sufficiently (i.e. only one tissue sample was subjected to this protocol), but extract 211 exhibited a good PCR performance. Protocol (A) is superior because it produces arDNA extracts of a good yield and with a low level of a PCR inhibition, that is, with a better consistency of PCR amplifications and applicability to different PCR marker systems and primers. Formalin-Steedman-arDNA usually did not require additional purification and/or diluting of arDNA extracts in order to be usable for PCR. Generally, the reproducibility and consistency of RAPD-PCR results were better with arDNAs produced by protocol (A) than with arDNA extracts produced by other tested protocols.

The most unsuitable protocols proved to be phenol-based protocols, especially for RAPD-PCR application. Experimental results from this study suggest that a phenol-based protocol is an unpredictable method for extracting PCR amplifiable arDNA. Tested phenol protocols gave relatively good yields, but most PCRs were inhibited. The phenol-based protocols also exhibited a high sensitivity to the smallest changes in extraction procedures that can affect the recovery of amplifiable DNA from preserved tissue. Protocol (C) of Shedlock *et al.* (1997) was the least useful phenol/chloroform protocol in this study ($p< 0.001$; Table 4). This protocol produced relatively good yields of formalin-arDNA from *Nezumia* specimens, but arDNA extracts were hardly applicable to any of the tested PCR marker systems (except for a few mitochondrial PCRs if a modified protocol (C) and diluted arDNA extracts were applied). Protocol (H) produced one ethanol-arDNA extract that was PCR non-amplifiable with any of the tested molecular markers and one of them (with applied overnight DNA precipitation) needed to be diluted in order to become PCR amplifiable. Three other ethanol-extracts (with DNA precipitation of 45 minutes) gave good PCR performances. The most successful phenol/chloroform protocol with formalin-*Nezumia* specimens was protocol ($G_1$), but mostly for mitochondrial markers (see Table 3).

**Table 6.** - *The arDNA extractions produced by different protocols in the study*

| Protocol | Species | No. DNA extr. | DNA concentrations (µg/ml) | | | | Suitability for PCRs in general | |
|---|---|---|---|---|---|---|---|---|
| | | | $n_1$ | Range | Mean | SE | Original | Modified |
| (A) | Formalin N.a | 45 | 35 | 1 - (over) 20 | 6.70 | 1.02 | good | |
| | Formalin N.m | 29 | 20 | 1 - (over) 20 | 10.03 | 1.82 | good | |
| | ***Formalin N.m + N.a*** | ***74*** | ***55*** | ***1 - (over) 20*** | ***7.91*** | ***0.94*** | | |
| | DMSO N.a (Greenl.) | 2 | 2 | 5 - 10 | 7.50 | 2.50 | very good | |
| | Frozen COD | 3 | 3 | 5 - (over) 20 | 15.00 | 5.00 | good | very good |
| | Fresh/frozen trout | 3 | 2 | 1.5 | 1.50 | 0.00 | good | very good |
| | 95% ethanol C.a | 2 | 2 | 20 - (over) 20 | (over) 20 | 0.00 | very good | very good |
| | ***Control specimens*** | ***10*** | ***9*** | ***1.5 - (over) 20*** | ***11.44*** | ***2.83*** | | |
| **Total (A):** | **All specimens** | **84** | **64** | **1 - (over) 20** | **8.41** | **0.91** | | |
| (B) | Formalin N.a | 83 | 46 | 1 - 4 | 1.18 | 0.09 | reasonable good | good |
| | Formalin N.m | 68 | 52 | 1 - 2 | 1.13 | 0.04 | reasonable good | good |
| | ***Formalin N.m + N.a*** | ***151*** | ***98*** | ***1 - 4*** | ***1.16*** | ***0.05*** | | |
| | Frozen COD | 8 | 5 | (over) 20 | (over) 20 | 0.00 | very good | very good |
| **Total (B):** | **All specimens** | **159** | **103** | **1 - (over) 20** | **2.10** | **0.40** | | |
| (C) | Formalin N.a | 14 | 13 | 1 - 20 | 5.35 | 0.99 | poor/ or inhibition | moderate/ or inhibition |
| | Formalin N.m | 17 | 16 | 1 - 20 | 9.53 | 1.65 | poor/ or inhibition | moderate/ or inhibition |
| | ***Formalin N.m + N.a*** | ***31*** | ***29*** | ***1 - 20*** | ***7.66*** | ***1.07*** | | |
| | Frozen COD | 11 | 10 | 15 - (over) 20 | 18.50 | 0.76 | good/ or inhibition | good/ or inhibition |
| **Total (C):** | **All specimens** | **42** | **39** | **1 - (over) 20** | **10.44** | **1.12** | | |
| (D) | Formalin N.a | 5 | 5 | 1 - 2 | 1.40 | 0.10 | moderate | moderate/ or good |
| | Formalin N.m | 10 | 4 | 1 - 2 | 1.50 | 0.20 | moderate | moderate/ or good |
| | ***Formalin N.m + N.a*** | ***15*** | ***9*** | ***1 - 2*** | ***1.44*** | ***0.10*** | | |
| | Frozen COD | 5 | 5 | 3 - (over) 20 | 16.60 | 3.40 | good | good |
| **Total (D):** | **All specimens** | **20** | **14** | **1 - (over) 20** | **6.86** | **2.31** | | |
| (F) | Formalin N.a | 1 | 1 | 8 | 8.00 | n/a | | very good |
| **Total (F):** | | **1** | **1** | **8** | **8.00** | **n/a** | | |
| ($G_1$) | Formalin N.a | 2 | 0 | | | | | moderate |
| | Formalin N.m | 2 | 0 | | | | | moderate |
| **Total ($G_1$):** | | **4** | **0** | | | | | |
| ($G_2$) | Formalin N.a | 4 | 0 | | | | | poor/ or moderate |
| | Formalin N.m | 4 | 0 | | | | | poor/ or moderate |
| **Total ($G_2$):** | | **8** | **0** | | | | | |
| ($G_3$) | Formalin N.m | 1 | 0 | | | | poor/ or moderate | |
| **Total ($G_3$):** | | **1** | **0** | | | | | |
| (H) | 95% ethanol N.a | 2 | 2 | 2 - (over) 20 | 11.00 | 9.00 | very good/ or inhibition | |
| | 70% ethanol N.a | 1 | 1 | 2 | 2.00 | n/a | very good | |
| | 95% ethanol N.m | 2 | 2 | 7 - 15 | 11.00 | 4.00 | very good/ or inhibition | |
| **Total (H):** | | **5** | **5** | **2 - (over) 20** | **9.20** | **3.60** | | |

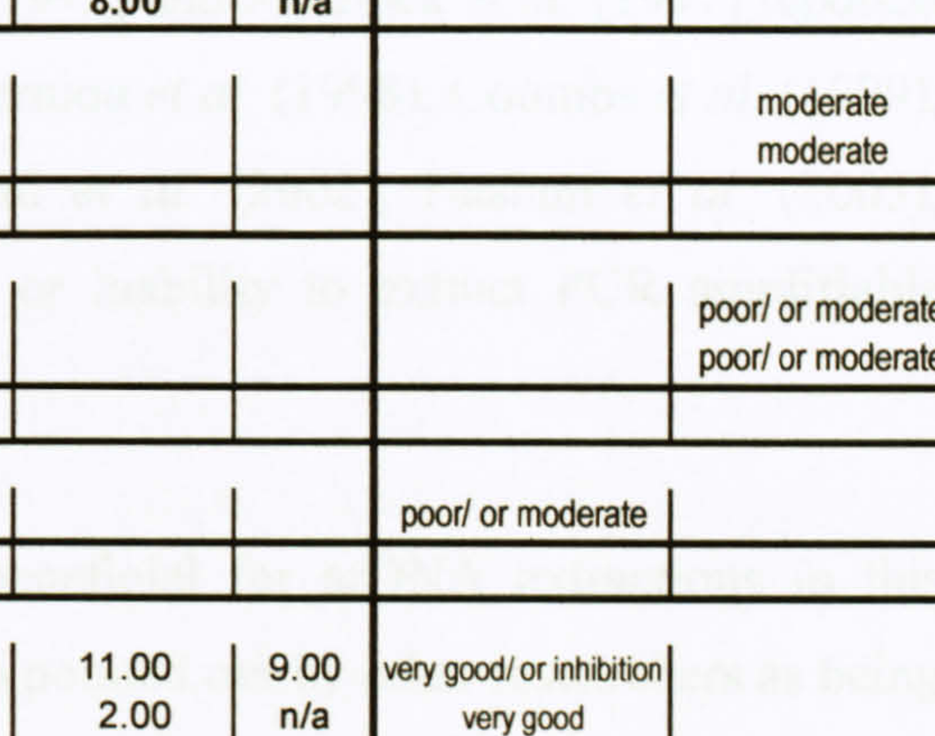
**$n_1$** Number of DNA extracts that DNA concentration was measured and recorded by taking a photograph
**N.m** *Nezumia micronychodon*
**N.a** *Nezumia aequalis*
**C.a** *Coryphaenoides armatus*
**SE** Standard error
**n/a** Not applicable

☐ Formalin *Nezumia aequalis* (N.a)

☐ Formalin *N. micronychodon* (N.m)

***Formalin N.m + N.a*** - Measurements for formalin-fixed specimens of *N. micronychodon* and *N. aequalis* together

***Control specimens*** - Measurements for control specimens together



Traditionally, PCI- (phenol : chloroform : isoamyl alcohol) based protocols include a lot of chemicals (e.g., phenol, SDS, NaCl, EDTA, sodium acetate) known to be strong PCR inhibitors and which potentially could cause problems in the application of such extracted DNA to the PCR (Weissensteiner and Lanchbury 1996; Cattaneo *et al.* 1997; Qiagen 2002; McNevin *et al.* 2005). In this study, the main problem with phenol protocols was indeed the presence of PCR inhibitors in DNA extracts that caused a failure of PCR amplifications. Montiel *et al.* (1997) and Mulligan (2005) suggest that PCI extraction typically fails to remove all PCR inhibitors from formalin-arDNA extracts (Mulligan 2005) and ancient DNA extracts (Montiel *et al.* 1997). Also, there is information about possible interactions between formalin and particular chemicals used in PCI-based protocols that might have negative effects both on DNA extraction and PCR amplification: for instance, formaldehyde and phenol (see Fig. 12; p 118), SDS and formalin (Jackson *et al.* 1990), formalin and proteinase K and chloroform-phenol solution (from: Eckerman 2006), as well as possible loss of DNA cross-linked with proteins if a phenol-chloroform extraction procedure is applied to formalin-preserved samples (Hasbun *et al.* 2005). These are the probable reasons for such poor results and inconsistency in PCR amplifications in this study, as well as contradictory results published on the usage of phenol-based protocols on formalin-fixed and other difficult specimens. For example, Shiozawa *et al.* (1992) and Shedlock *et al.* (1997) reported good yield of PCR amplifiable arDNA, whereas Herniou *et al.* (1998), Coombs *et al.* (1999), Sato *et al.* (2001), Godhe *et al.* (2002), Legrand *et al.* (2002), Hasbun *et al.* (2005), Chakraborty *et al.* (2006) experienced problems or inability to extract PCR amplifiable arDNA if phenol-based protocols were applied.

Conversely, guanidinium-based protocols were beneficial for arDNA extractions in this study. Guanidinium-thiocyanate (GuSCN) was also pointed out by other researchers as being a useful chemical in DNA extractions from museum and other difficult specimens (Hammond *et al.* 1996; Fredricks and Relman 1998; Whittier *et al.* 1999; Konomi *et al.* 2002; Umetsu *et al.* 2002; Rohland *et al.* 2004).

The results from this study (on investigated preserved specimens) strongly suggest that the yield of arDNA and quality of arDNA extracts, i.e. PCR usability of arDNA extracts, depends on the DNA extraction protocol applied, including pre-treatment techniques (washing/drying of preserved tissue) prior to the extraction of DNA. Also, it was acknowledged that it is much more important to obtain arDNA extracts free of PCR inhibitors than arDNA extracts with higher DNA concentrations.

Comparative experiments with formalin-fixed, Steedman's-preserved and control DMSO- and ethanol-preserved specimens of *Nezumia* strongly indicate the significance of the preservation regime for the recovery of DNA from preserved fish specimens. Additional evidence on mackerel further supports this statement (see sections 3.3.1. and 3.4.2.).

#### *3.1.1.3. Type of tissue and size of tissue used in extracting formalin/Steedman-arDNA*

Muscle tissue, frequently used for "standard" DNA studies in fish, proved to be a good tissue type for DNA extraction from preserved specimens (see Appendices 1 – 4 and Fig. 13-Lanes: 6-11). Mixed tissue samples (lower and upper parts of small fish with all internal organs, but with removed skin and intestines because of the possible presence of contaminant DNA from parasites and/or partly digested food material) gave high DNA concentrations in the archival DNA extracts (Appendix 4 – Spots: I-5, III-2, III-3, IV-6 and V-6 with DNA concentrations between 8 μg/ml and 20 μg/ml) and good PCR performances as well (see the evidence on support material on CD). Eye, spine (with and without vertebrae) and brain tissue seem to be very good sources of arDNA (Appendix 2 – Spots: I-5, 6; Appendix 4 – Spots: V-1, 4; III-2, III-3, with DNA concentrations between 10 μg/ml and 20 μg/ml; see also support material on the CD for PCR performances). Liver was investigated, but results were inconsistent. Sometimes arDNA extracted from liver gave very good PCR results, especially if a DNA pellet was subjected to prolonged washing with 70% ethanol, but a few formalin-arDNA extracts from liver caused a complete PCR inhibition (a good yield of DNA was obtained, but probably with a high presence of inhibitors that caused difficulties in using this arDNA for the PCR; see support material on the CD). A few attempts were made to extract DNA from bones, but most extracts showed a low yield of non-amplifiable DNA.

The amounts of tissue necessary for successful arDNA extraction also were observed in the project. The required amount of tissue depended on the type of tissue used. If muscle tissue was used, 100-300 mg of wet tissue (0.2 $cm^3$ – 0.5 $cm^3$) was usually sufficient for producing one arDNA extract of 40-50 μl with a DNA concentration between 1 μg/ml and 15 μg/ml, although 20 μg/ml was also achievable by using protocol (A). If spine, kidney, or brain was used, 20-100 mg of wet tissue was a sufficient amount.

### *3.1.1.4. A possible suitability of particular extraction protocols for particular PCR marker systems and primers*

From results of this study, there are indications that particular DNA extraction protocols and/or washing/drying regimes were more suitable to amplify a particular gene sequence than others.

An example of an arDNA extract produced by a particular protocol that might be more suitable for one particular PCR marker system, but not for another (*Nezumia*-specific, RAPD-derived and mitochondrial in this case) is arDNA extract no. 215 produced by guanidinium-based protocol (A). This arDNA extract generated a good PCR product with the *Nezumia*-specific primer set 18 (Fig. 16, gel (b)-Lane 9 [p 126]), but a very faint band (poor PCR yield) with the mitochondrial 16S primer set (Fig. 16, gel (b)-Lane 3 [p 126]). This might suggest damage of the mitochondrial genome in a particular fish individual due to preservation, but it also might be related to the extraction protocol applied, i.e. a higher possibility of recovering particular genome regions if a specific DNA extraction protocol is applied (see also supplementary evidence on mackerel; sections 3.3.1. and 3.4.2.).

Another example for a possible effect of DNA extraction protocol on arDNA being amplifiable with a particular PCR marker system but not with another (RAPD and mitochondrial in this case) are arDNA extracts no. $222^{F}$ (produced by phenol-based protocol ($G_I$)) and arDNA extract no. 85(II) (produced by guanidinium-based protocol (A)) from the same fish individuals (fish no. 3; R-3). ArDNA extract no. $222^{F}$ was unsuccessful in RAPD-PCR amplifications (with RAPD primer 48: Fig. 19, gel (a)-Lane: 15, and with RAPD primer 44: Fig. 19, gel (b)-Lane: 15; p 130), but it was successful in amplifying the mitochondrial 16S gene region (Fig. 19, gel (c)-Lane: 15). Conversely, arDNA extract no. 85(II) produced by protocol (A) was successful in RAPD-PCR amplifications with both RAPD primers (Fig. 19, gels (a) and (b)-Lanes 13 in both gels), but unsuccessful in amplifying the mitochondrial 16S gene region (Fig. 19, gel (c)-Lane: 13).

An example for a possible effect of DNA extraction protocol on arDNA being amplifiable not only with a particular PCR marker system (RAPD and mitochondrial in this case), but also with sequences from a particular genome region, i.e. primers (RAPD primer 44 and RAPD primer 48 in this case), is comparison arDNA extracts produced by two guanidinium-

based protocols ((A) and (B)): extract no. 213 produced by protocol (A), and extract no. 137 produced by protocol (B) (see Figure 19 [p 130]; gels (a),(b),(c)-Lanes 10 for extract 213 and gels (a),(b),(c)-Lanes 8 for extract 137). The arDNA extract no. 213 generated PCR products with RAPD primer 44 (Fig. 19; gel (b)-Lane: 10), but not with RAPD primer 48 (Fig. 19; gel (a)-Lane: 10), and not with mitochondrial primers (Fig. 19; gel (c)-Lane: 10). However, arDNA extract no. 137 gave very successful PCR amplifications with both RAPD primers (Fig. 19; gel (a)-Lane 8 for RAPD primer 48 and gel (b)-Lane 8 for RAPD primer 44), but not with mitochondrial primer set 16 (Fig. 19 gel (c)-Lane: 8). None of these two arDNA extracts produced by guanidinium-based protocols amplified 16S mitochondrial sequence (Fig. 19; gel (c)-Lanes: 8, 10). These two arDNA extracts (nos. 137 and 213) were from the same species (*Nezumia micronychodon*) but from two different fish individuals. This might suggest that different DNA extraction protocols (in this case protocol (A) for arDNA extract no. 213 [muscle tissue], and protocol (B) for arDNA extract no. 137 [vertebra with spinal cord]) might be more suitable for a particular RAPD primer. But, this also might suggest different degrees of sequence damage in different fish individuals, or in different types of tissue.

An example for a possible effect of DNA pre-extraction treatment on arDNA being preferably amplifiable with a particular PCR marker systems but not with another (RAPD and mitochondrial in this case) is arDNA extract no. 206 produced by guanidinium-based protocol (A). For this arDNA extract, only drying of the preserved tissue sample at 37°C for a few days (without any tissue pre-washing in $sdH_2O$ and/or in 1xGTE buffer) was applied as a pre-extraction treatment. This arDNA was amplifiable for the mitochondrial 16S region (Figure 19 [p 130]; gel (c)-Lane 6), but not for RAPD-PCRs (Figure 19: gel (a)-Lane 6 for RAPD primer 48 and gel (b)-Lane 6 for RAPD primer 44). The data from this study suggest that formalin-arDNA extracts produced by guanidinium-based protocol were usually better amplifiable with RAPD PCR marker system than with mitochondrial (see also support material on the CD), but in this particular experiment it was the opposite. This might suggest that a particular pre-extraction treatment (in this case drying a preserved tissue sample at 37°C without any pre-washing) produced arDNA extracts that were more applicable to one (or to some) particular PCR marker system(s), perhaps because of better recovery of arDNA from a particular region(s) of the genome. More examples of PCRs that were performed with arDNA extracts for which tissue samples were subjected to different pre-extraction treatments are presented in support material on the CD.

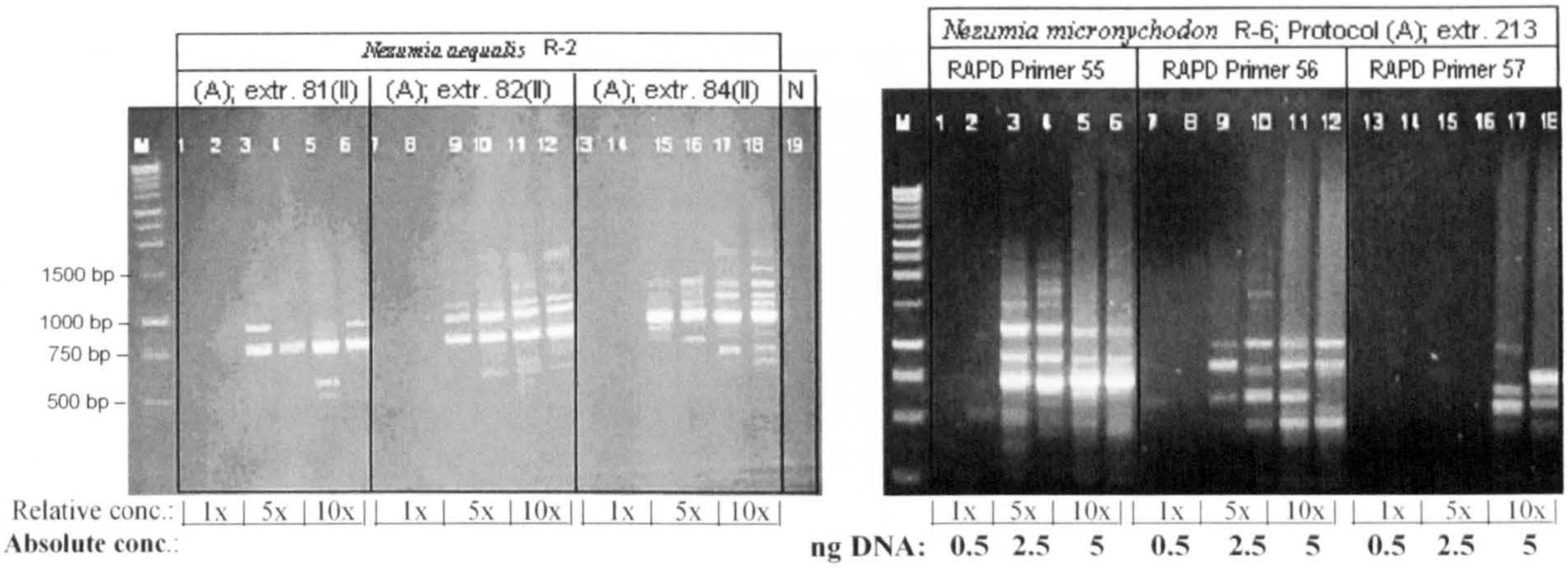
Considering the fact that different protocols involve different chemicals, reagents and treatments, differences in quality and yield of DNA are likely to occur. It is more likely that differences are expressed between nuclear and mitochondrial DNA because of, for example, different binding of proteins with nDNA and mtDNA that may affect their ease of extraction and intactness. Possible interactions (unknown to date) of chemicals used for tissue/organism preservation and reagents/chemicals used for DNA extractions are additional complicating factors which might be important in the accessibility and recovery of DNA from preserved organisms (i.e. particular regions and sequences of genomes). The tissue type and cell environment for DNA recovery from preserved specimens are also important factors that need to be considered. Experimental data from this study suggest that some extraction protocols and, perhaps, washing/drying a tissue sample prior to DNA extraction, are better in recovering particular regions of genome (DNA sequences), but this requires much more extensive research and much better understanding of the nature of arDNA (in differently preserved specimens of different organisms/species), i.e. much more evidence before it is possible to draw a general, concrete conclusion about this.

### 3.1.2. Developing approaches for RAPD-PCR investigations with formalin-arDNA

RAPD-PCR methodology requires reproducible PCR results in order to carry out reliable and accurate analyses. Many factors (the quality of a PCR template, differences between DNA preparations, the presence of PCR inhibitors in a DNA extract, etc.) could affect the reproducibility of the RAPD pattern (Micheli *et al.* 1994). This study attempted to optimise the RAPD-PCR methodology for its traditional use on formalin-fixed, Steedman's preserved specimens of fish.

First, one RAPD primer was tested against different arDNA extracts (produced by the same extraction protocol) from one fish individual (Fig. 21, gel (a)), and then, different RAPD primers were tested against one formalin-Steedman's-arDNA extract (Fig. 21, gel (b)). A primer was tested against three different PCR-DNA template concentrations (relative PCR-DNA template concentrations; **1x**: 0.5 µl of arDNA suspension in 25 µl of PCR reaction, **5x**: 2.5 µl of arDNA suspension in 25 µl of PCR reaction, and **10x**: 5 µl of arDNA suspension in 25 µl of PCR reaction). The reproducibility of some RAPD fragments (e.g. for sizes of 750 bp and ~900 bp; Fig. 21, gel (a)-Lanes: 3-6, 9-12 and 15-18) was achieved to some extent

with 5x and 10x relative concentrations of arDNA, but insufficiently for an adequate application of the RAPD-PCR methodology.




**(a) – RAPD primer 44** **(b) – three different RAPD primers**

**Figure 21** – ***Developing a reliable approach for RAPD-PCR experiments.*** The importance of performing RAPD-PCRs with, at least, three different PCR-DNA template concentrations in duplicates: **1x** (0.5 µl DNA in 25 µl PCR reaction volume), **5x** (2.5 µl DNA in 25 µl PCR reaction volume) and **10x** (5 µl DNA in 25 µl PCR reaction volume).
**Gel (a)** represents RAPD-PCR results with the same RAPD primer (Primer 44), but with three different arDNA extractions of one *Nezumia aequalis* fish individual.
**Gel (b)** represents RAPD-PCRs with the same arDNA extract (no. 213 - *Nezumia micronychodon*), but using three different RAPD primers.
**R** – Identifier of fish individual, i.e. a fish specimen (R-2 is fish no. 2; R-6 is fish no. 6)
**(A)** – DNA extracts produced by protocol (A)
"M" indicates size marker fragments - 1 Kb DNA Ladder (Promega).
"N" indicates negative control of PCR reaction (sterile water was used as PCR template instead of DNA).

Attempts to put the same absolute concentration of DNA in RAPD-PCR reactions (for instance, 0.5, 5 or 10 ng in 25 µl of PCR reaction volume) did not solve the problem. For some arDNA extracts and RAPD primers this worked (see Fig. 21-gel (b) and Fig 22) in their duplicates, but it did not work for the majority of arDNA extracts and RAPD primers, or across different PCR-DNA template concentrations. For instance, it worked in PCR duplicates for arDNA extracts nos. 213 and 215 with RAPD primer 28 and with 5 ng DNA in PCR reaction (Fig. 22, Lanes: 11 and 12 in both gels). Also, it was applicable for RAPD primer 55 (Fig. 21, gel (b)-Lanes: 5 and 6). However, this strategy (to put the same amount of arDNA into PCR reactions) did not work, for instance, with RAPD primers 56 and 57 with the same arDNA extract 213 that was successfully applied for primers 28 and 55 (see Figs. 21 and 22). This might be related to the robustness of a particular RAPD primer, but not necessarily. For example, these two arDNA extracts, nos. 213 and 215 (both produced by

the same extraction protocol - protocol (A) and from the same species (*Nezumia micronychodon*), but from two different fish individuals) gave very different performances (related to the reproducibility of the RAPD patterns and RAPD fragments in duplicates and across different PCR-DNA template concentrations and PCR experiments) with two RAPD primers (RAPD primers: 28 and 29) - compare gel (a) and (b) in Figure 22. The data strongly suggest that the success of RAPD-PCR amplifications and the reproducibility of the RAPD patterns (or particular RAPD fragments/bands) depend not only on the RAPD primer used and concentrations of DNA used as a PCR template, but even more on the particular arDNA extract (and, may be, fish individual). More evidence is provided in support material on the CD.

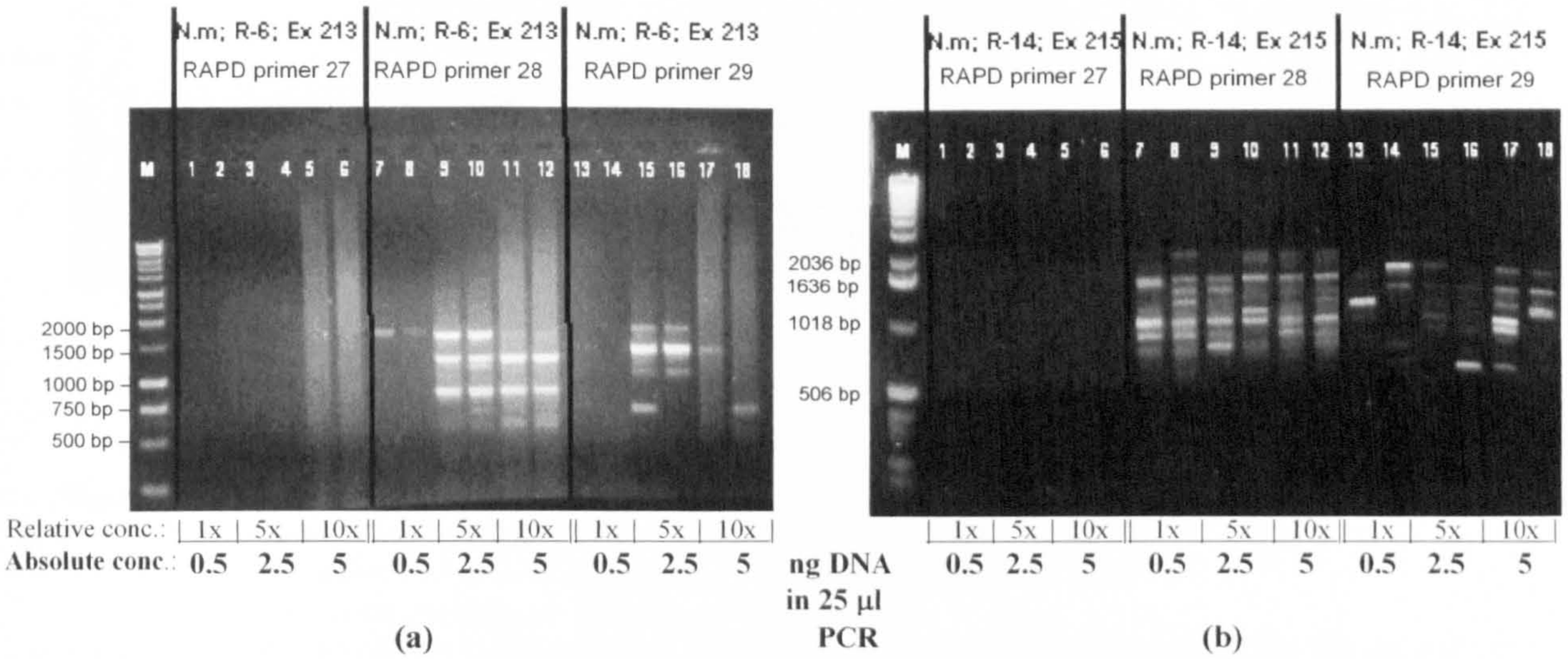



**Figure 22 – Testing three different RAPD primers** with the same concentrations of DNA in PCRs (two arDNA extractions of two fish individuals of *Nezumia micronychodon* (N.m) produced by protocol (A) – arDNA extract no. 213: gel (a); and arDNA extract no. 215: gel (b)).
**R** – Identifier of fish individual, i.e. a fish specimen (R-6 is fish no. 6; R-14 is fish no. 14)

The "behaviour" of all RAPD-fragments from the profiles, related to their yield and PCR-DNA template concentrations used, was not equal. For instance, by increasing the concentration/amount of DNA as a template in PCR, some bands gained intensity (better yield; Fig. 21(b)-compare lanes 3 and 4 with lanes 5 and 6 for a band size of ~750 bp), whereas others faded (Fig. 21(b)-compare lanes 3 and 4 with lanes 5 and 6 for a band size of ~1200 bp), or even disappeared from the profile (Fig. 22(a) – compare lanes 9 and 10 with lanes 11 and 12 for a band size of ~1700 bp, or lanes 15 and 16 with lanes 17 and 18 for bands sizes of ~1500 bp and ~1800 bp).

Because of the indications that the reproducibility of RAPD results might be related to a particular fish individual, as well as to a particular RAPD primer and arDNA extract, further testing involved different fish individuals with one primer (see, for example, Figs. 23 and 24 with RAPD primer 37), and then with different primers (see support material on the CD).

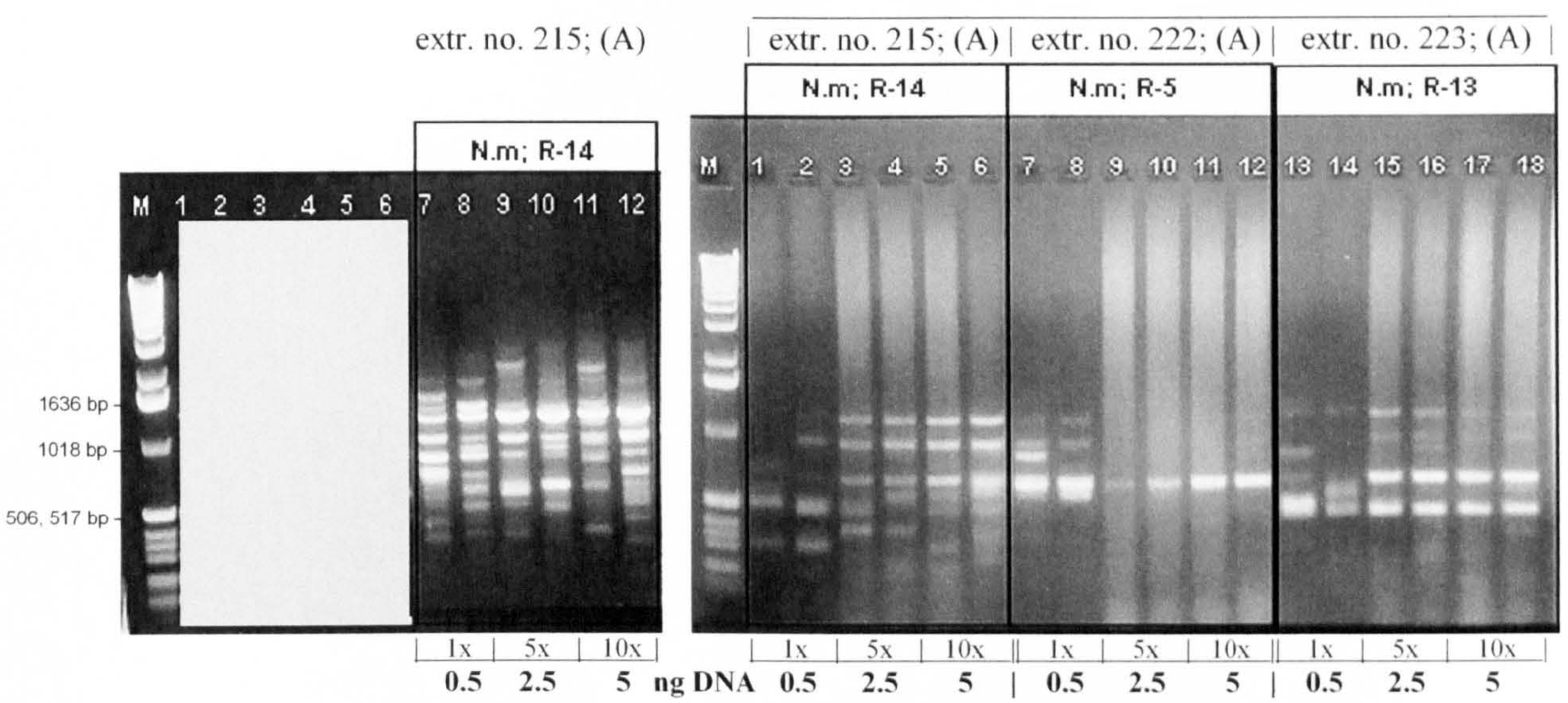



**Figure 23** – RAPD-PCR experiments with formalin-arDNAs of *N. micronychodon* (N.m) (different arDNA extracts and from different fish individuals) using RAPD Primer 37.
"M" indicates size marker fragments - 1 Kb DNA Ladder (Gibco).
**(A)** – DNA extracts produced by protocol (A)
**R** – Identifier of fish individual, i.e. a fish specimen (R-5 is fish no. 5; R-13 is fish no. 13; R-14 is fish no. 14)

Differences were recorded in the RAPD-PCR banding patterns of the same individuals (individual fish specimen) if different arDNA extracts (produced by the same DNA extraction protocol) were applied (Fig. 21 (a) for RAPD primer 44), and even with the same arDNA extracts run in separate PCR experiments (for example, extract no. 215; Fig. 23; compare gel (a)-Lanes: 7-12 with gel (b)-Lanes: 1-6 for RAPD primer 37). Differences in RAPD-PCR results were also observed with the same arDNA extract but applying different PCR-DNA template concentrations (Figs. 21 - 24). Differences in RAPD-PCR profiles (the number of bands, sizes and the intensity of bands) would occur even in PCR reaction-duplicates (PCR reactions with the same content) that were run simultaneously (Figs. 21-24; for example in Fig. 24; compare in gel (b) lanes 6 and 7, 8 and 9, 10 and 11 for formalin-arDNA extract no. 210, or compare in gel (d) lanes 7 and 8, 9 and 10, 11 and 12 for extract no. 241).

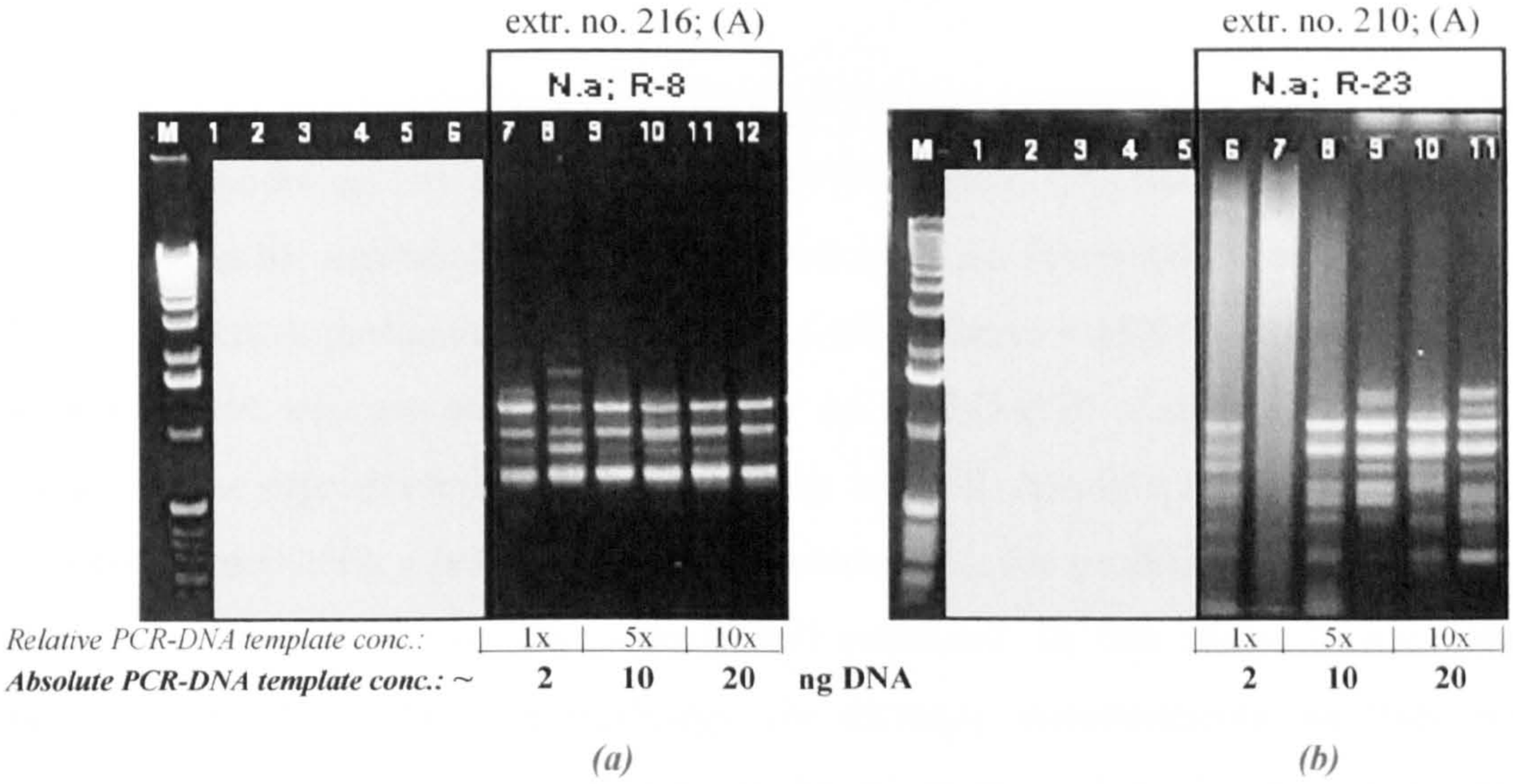



**Figure 24** – RAPD-PCR experiments with formalin-arDNA of *N. aequalis* (N.a) (different arDNA extracts from different fish individuals) using RAPD Primer 37.
"M" indicates size marker fragments - 1 Kb DNA Ladder (Gibco).

All these tests were carried out in order to determine the sensitivity and applicability of the RAPD methodology on investigated preserved specimens; that is, in order to optimise RAPD-PCRs for investigated samples (if possible). As described in section 3.1.1., different DNA extraction protocols produced formalin-Steedman's-arDNA extracts that performed with different success and reproducibility of RAPD-PCR results (see Fig. 19; p 130). Although the reproducibility and consistency of PCR results were increased by producing arDNA extracts with a newly developed protocol (A), the problem of full reproducibility of RAPD-PCR results across experiments still remained. In this study, it was attempted to optimise the RAPD-PCR methodology for distance measurements for two investigated species of *Nezumia* on formalin-fixed, Steedman's preserved specimens of fish, but this goal could not be achieved with the investigated preserved specimens, although some RAPDs were reproducible in the majority of PCR amplifications (see support material on the CD). Other researchers have also tried to use the RAPD methodology on formalin-fixed samples, but their success was variable – from the lack of successful RAPD-PCR amplifications (Gurdebeke and Maelfait 2002) to generating partially reproducible band patterns (Eckerman and Welsh 1997; Siwoski *et al.* 2002).

The low and degraded DNA content of formalin-fixed samples makes PCR amplifications difficult to optimise, and there is potential for erroneous results when some loci fail to be amplified. The sizes of generated RAPD fragments depended on the RAPD primer and the particular arDNA extract, i.e. relative PCR-DNA template concentrations (template volumes) applied in PCRs. Reproducible RAPD band sizes of over 1000 bp were generated with some primers suggesting that formalin-arDNA is not so fragmented in some regions of genome, but their reproducibility (especially related to their yield, i.e. the intensity of bands) was to some extent lower than with RAPD fragments sizes smaller than 1000 bp (Figs. 21-24; see also support material on CD). The possibility of "PCR competition" (Hallden *et al.* 1996) with fragmented and degraded DNA (as arDNA is) might significantly increase differences in RAPD-PCR results. If a few priming sites are present in the investigated genome sequence with a particular RAPD primer, it is expected that smaller RAPD fragments will be in favour of PCR amplification, except if these targeted sequences are damaged, or if DNA sequences (or parts of a sequence) are masked by protein (due to crosslinking), making these sequence regions unusable for PCR (i.e. thermostable polymerase).

### *3.1.2.1. Possible causes and explanations of the RAPD-PCR inconsistency with arDNA*

Inconsistent results were partly because of the sensitivity and known pitfalls of the RAPD-PCR technology in general (see section 1.3.5.1. in General Introduction), but the quality of the arDNA (damage due to formalin/Steedman's preservation) and the presence of diffusible PCR inhibitors in arDNA extracts are obviously the major problems that caused irreproducibility of the RAPD patterns and/or particular RAPD fragments (bands). It was not expected to find a very high level of the RAPD-PCR reproducibility in these experiments, but it was unexpected to find such a high level of the inconsistency and irreproducibility across different concentrations of the same PCR-DNA template (DNA of one arDNA extract), and especially in the duplicates of PCRs (the same content of two PCR reactions prepared from the same PCR master-mix and which were run in parallel next to each another in the same PCR machine).

The sensitivity of RAPD-PCR technology to the smallest changes of PCR conditions and/or set ups of PCR reactions add complexity in generating reproducible RAPD results in PCR experiments. For instance, small errors might occur in pipetting DNA in duplicates, especially with small volumes (e.g., with 0.5 µl of PCR-DNA templates). Small pipetting errors are mostly insignificant in PCRs with good DNA, but it could be of a high significance if PCRs involve degraded and fragmented DNA combined with a variety of PCR inhibitors in DNA extracts. Furthermore, the smallest differences in temperatures in the PCR thermal cycler across a heating block might also affect the reproducibility of RAPD-PCR results.

Considering the facts that during experiments in this study only a high quality of pipette tips and pipettors (which were regularly calibrated) were used and pipetting was performed carefully, PCR machines were regularly checked and serviced and, as a precaution, the outside rows (spaces) of the PCR block were always left empty in case of a possible influence of ambient temperature on the temperature in the heating block (this, probably, was unnecessary because today's thermal cyclers are made with good thermal isolation), that all arDNA suspensions in storage tubes were well-mixed before pipetting for adding DNA into PCR reactions/mastermix, all tubes with PCR reagents and arDNAs were kept on ice during PCR set ups, and all PCR reactions were prepared from the same mastermix for one PCR experiment (containing all PCR components except DNA and/or primer), it is difficult to explain a complete amplification failure in one of the PCR reaction-duplicates (e.g., Fig.

24, gel (b)- compare lanes 6 and 7, or gel (d)-lanes 11 and 12, and lanes 13 and 14) by technical deficiency in setting up and performing RAPD-PCR amplifications. This occurred regardless of the amounts and concentrations of DNA in RAPD-PCR reactions. Also, it is difficult to explain such big differences in the RAPD-PCR banding patterns; that is, differences in the intensity and number of bands in RAPD profiles in duplicates (e.g., Fig. 24, gel (b)-Lanes:10 and 11; or gel (d) - PCR-duplicates in all RAPD-PCRs) just with the known pitfalls of RAPD-PCR methodology. It seems that causes of the RAPD`s "erratic behaviour" with archival DNA (arDNA) are much more complex.

One explanation might be that the number of DNA molecules with sequences that contain different kinds of damage (cross-links, nicks, deletions, base modifications) and amounts of diffusible PCR inhibitors are unequally distributed in different arDNA extracts and maybe even in PCR-DNA templates of one arDNA extract. The quality of an arDNA extract depends on how many cells were included in DNA preparations that consisted of seriously damaged, moderately damaged and undamaged genome sequence (i.e. targeted regions of sequences). Adding the presence of diffusible PCR inhibitors (of unknown kinds and amounts), all together could make very heterogeneous DNA samples (arDNA extracts and PCR-DNA templates). The reproducibility of RAPD-PCR results depends very much on the quality and homogeneity of arDNA extracts and PCR-DNA templates, as well as on the ability of *Taq* polymerase to access targeted genome sequences and continue polymerisation by correctly incorporating bases for DNA damaged regions.

In my opinion, the heterogeneities of arDNA extracts and PCR-DNA templates, i.e. differences in the number of PCR amplifiable DNA sequences and the presence of "diffusible" PCR inhibitors (chemicals that co-purified with DNA and act as PCR inhibitors) in the PCR-DNA templates, are the main problems and the most probable reasons of a such inconsistency in RAPD-PCR results (see hypothetical model in Fig. 25), as well as in the conjunction with the "PCR competition" that often occur in RAPD-PCRs (Hallden *et al.* 1996).

Hypothetically, DNA molecules (genome sequences) that come from different cells (of a formalin-preserved tissue) might have intact DNA sequences (the DNA sequence is undamaged and can be used in full length for *Taq* polymerase), or DNA sequences might be damaged to different extents: **point modifications** (missing and/or altered bases in the sequence at one or more locations along the DNA sequence) and/or **length modifications**

(the sequence damage spans larger parts of sequence at one or more locations: cross-links, deletions/insertions and/or distortions). Between two extreme points (DNA sequence being fully PCR amplifiable and DNA sequence being completely non-amplifiable [PCR unusable]), there are, probably, different degrees of damage to DNA sequences - some regions of DNA sequence are PCR amplifiable, but some are not. It might be possible to express the degree of their PCR usability in percentages of DNA sequence damages, i.e. how much of arDNA sequence from one cell (DNA molecule) is potentially amplifiable/non-amplifiable (10%, 30%, 50%, 70%, or 90%) - see Fig. 25.

The PCR usability of arDNA is complicated by the presence of different kinds of "diffusible" PCR inhibitors that also might come into arDNA extracts and PCR-DNA templates in different ratios. The nature of PCR inhibitors present in arDNA extracts is still unknown, but probably it is possible to categorise them in serial from very strong PCR inhibitors (labelled as 1) to very mild (labelled as 7) - see Fig. 25.

Applying the basic rule of probability in the case of a non-homogenous sample, i.e. a heterogeneous DNA extract, it is obvious that we cannot control and select how many DNA molecules with undamaged DNA sequences, or DNA molecules with specific degrees of sequence damage, will be taken by ordinary pipetting in 0.5 μl, 2.5 μl or 5 μl of PCR-DNA template and be included in a particular PCR reaction. Neither can we select how many molecules of strong or mild PCR inhibitors will "accompany" these DNA molecules in a particular PCR-DNA template. Because of so many unknowns about arDNAs and formalin/Steedman's-arDNA extracts (non-diffusible and diffusible PCR inhibitors), it is difficult to guess how heterogeneous arDNA samples really are - how many damage variations in DNA sequences exist and how many types of PCR inhibitors is it possible to expect in one arDNA extract produced by a specific DNA extraction protocol.

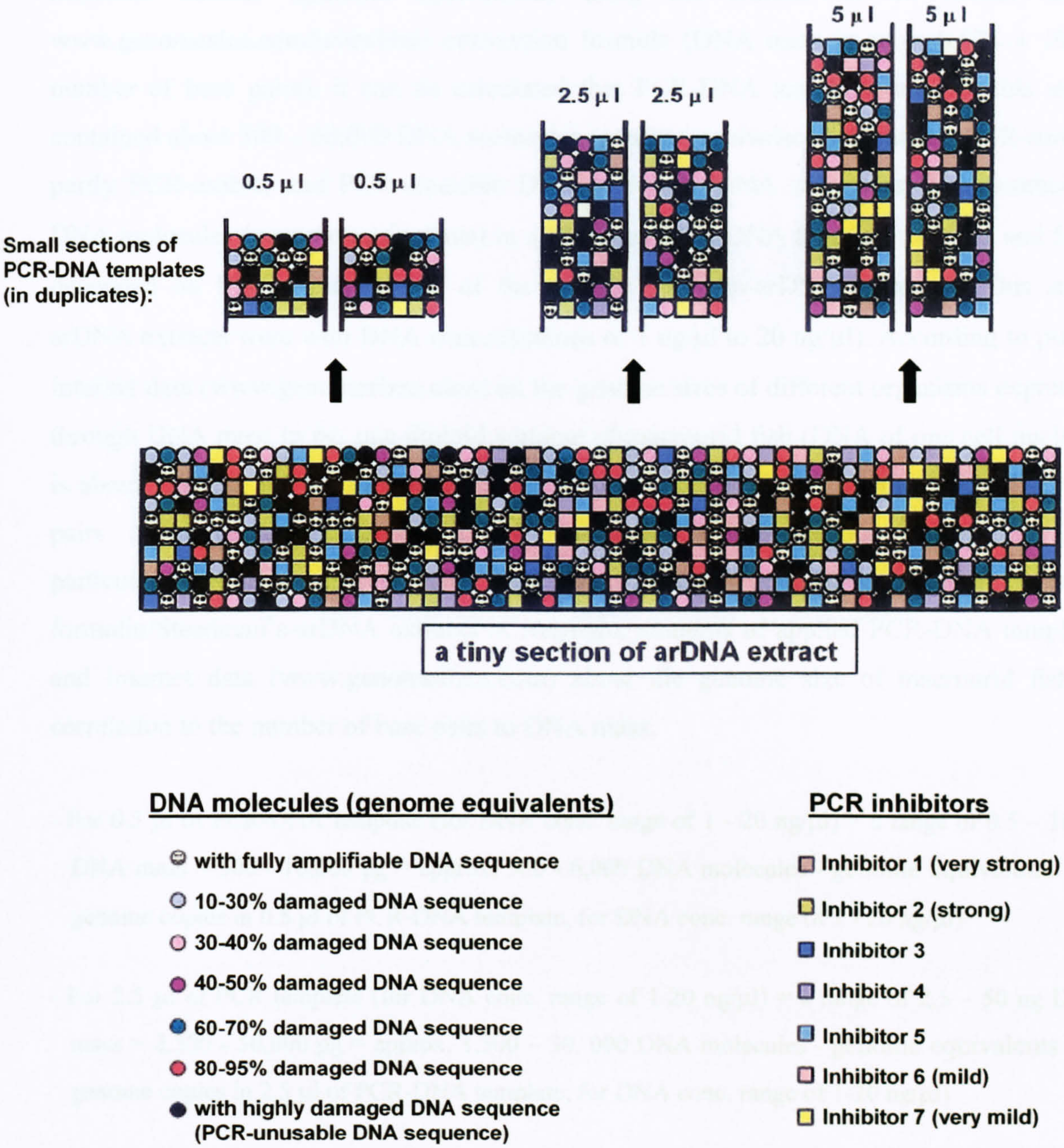



**Fig. 25** – Hypothetical, graphic view of the possible scenario that causes differences in RAPD-PCR products of PCRs with different PCR-DNA template concentrations and in duplicates with the same arDNA extract. This is a hypothetical example with 7 types of differently damaged and artificially altered DNA sequences, or otherwise PCR-unusable DNA molecules (at certain degree), and in combination with 7 different kinds of PCR inhibitors (i.e., compounds of arDNA extracts that have a different effect on the PCR inhibition – from very strong to very mild).

The question is: "How many DNA molecules (genome DNA sequences) are really PCR usable/amplifiable and how many different variants of PCR inhibitors are present in arDNA extracts/PCR-DNA template?" Applying basic and approximate calculations of the DNA molecule number (genome equivalents) using the Dolezel *et al.* (2003; from: www.genomesize.com/units.htm) conversion formula (DNA mass in pg x 0.978 x $10^9$ = number of base pairs), it can be calculated that PCR-DNA templates used in this study contained about 300 – 60,000 DNA molecules (genome equivalents), including PCR-usable, partly PCR-usable, and PCR-unusable DNA molecules/DNA sequences. The number of DNA molecules (genome equivalents) in a particular PCR-DNA template (0.5, 2.5 and 5 µl) depended on DNA concentration of the specific formalin-arDNA extract (in this study arDNA extracts were with DNA concentrations of 1 ng/µl to 20 ng/µl). According to public Internet data (www.genomesizes.com) on the genome sizes of different organisms expressed through DNA mass in pg, one diploid genome of macrourid fish (DNA of one cell nucleus) is about 1.64 pg, i.e. the diploid genome size of macrourid fish is about 1.64 billion base pairs. Approximate calculations of DNA molecule numbers (genome equivalents) in a particular PCR-DNA template have been based on DNA concentration measurement of formalin/Steedman's-arDNA extracts of *Nezumia*, amounts of applied PCR-DNA template, and Internet data (www.genomesizes.com) about the genome size of macrourid fish in correlation to the number of base pairs to DNA mass:

- For 0.5 µl of PCR-DNA template (for DNA conc. range of 1 - 20 ng/µl) = a range of 0.5 – 10 ng DNA mass = 500 - 10,000 pg = approx. 300 - 6,000 DNA molecules - genome equivalents (i.e. genome copies in 0.5 µl of PCR-DNA template, for DNA conc. range of 1 - 20 ng/µl)

- For 2.5 µl of PCR template (for DNA conc. range of 1-20 ng/µl) = a range of 2.5 – 50 ng DNA mass = 2,500 - 50,000 pg = approx. 1,500 – 30, 000 DNA molecules - genome equivalents (i.e. genome copies in 2.5 µl of PCR-DNA template, for DNA conc. range of 1-20 ng/µl)

- For 5 µl of PCR template (for DNA conc. range of 1-20 ng/µl) = a range of 5 – 100 ng DNA mass = 5, 000 - 100,000 pg = approx. 3,000 – 60, 000 DNA molecules - genome equivalents (i.e. genome copies in 5.0 µl of PCR-DNA template, for DNA conc. range of 1-20 ng/µl).

A genome sequence usually has a few priming sites for a particular RAPD primer – meaning that the above figures are multiplied with numbers of possible priming sites in the genome with a particular RAPD primer. However, these calculated figures of the total number of available DNA molecules/sequences in PCR templates contradict the PCR amplification

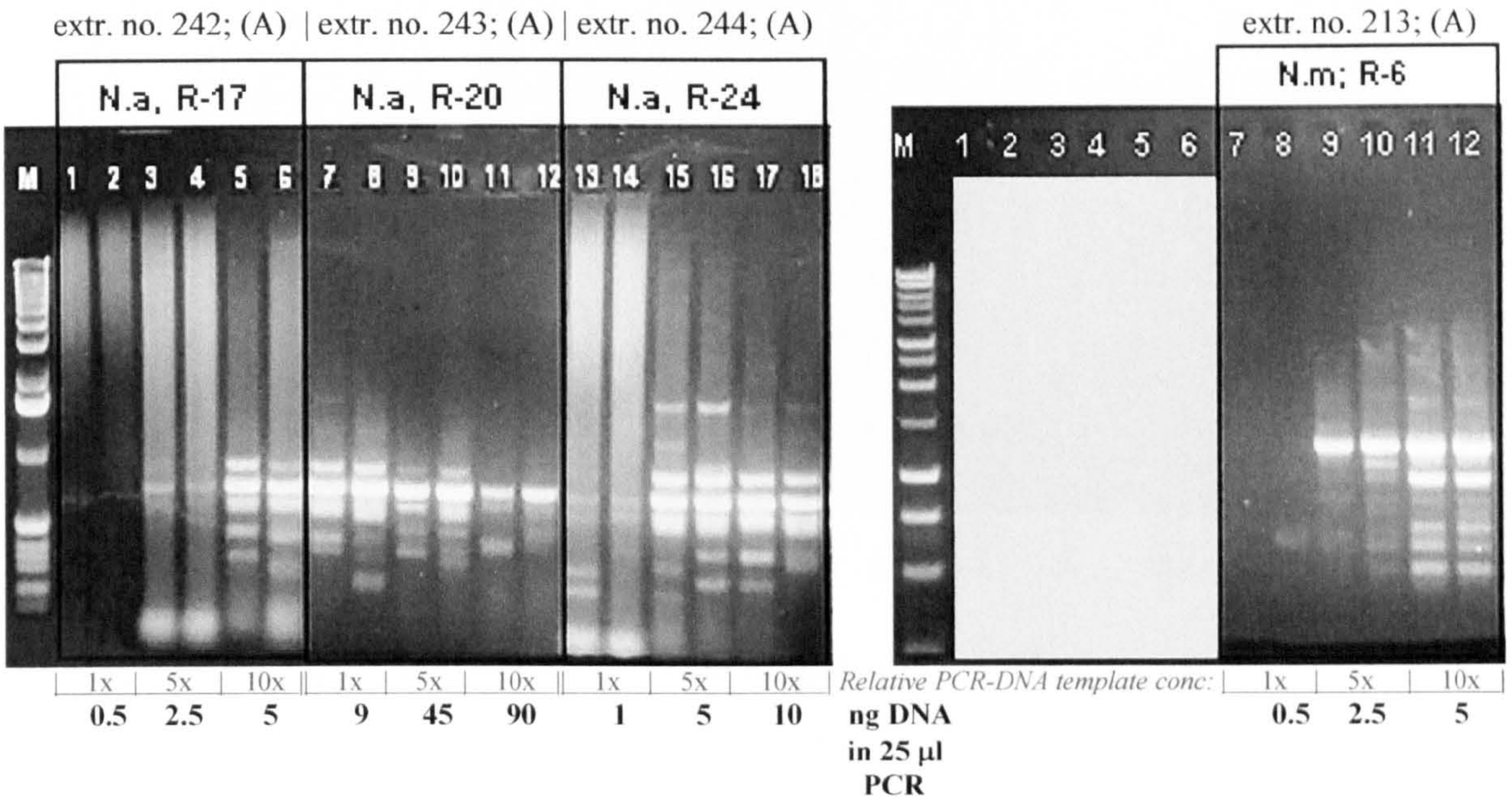
success rate in this study. Experimental data from the study suggest that a much lower number of PCR-usable DNA molecules/sequences from formalin-arDNA extracts might be available than theoretically expected from the above calculations. Or, another assumption might be that the PCR inhibition effect is much stronger than presumed, and probably depends on how many molecules of strong PCR inhibitor(s) are present in a particular PCR-DNA template. It seems that the variety of diffusible PCR inhibitors that come from preservation fluid and, probably, interactions of these chemicals and those used in DNA extraction procedures, are in much higher concentrations and with very different effects on PCR.

However, it is still unclear whether the major PCR difficulties occur because of "non-diffusible" PCR inhibitors (i.e., damage on DNA sequence: modifications, fragmentation, crosslinking), or because of the PCR inhibition caused by "diffusible" PCR inhibitors (chemicals that co-purified with arDNA and act as PCR inhibitors).

In my opinion, based on evidence from this study, it seems that the presence of diffusible PCR inhibitors has more adverse effect on the PCRs than low availability of amplifiable arDNA. Experimental data from this study have confirmed the significance of reducing the concentration of "diffusible" (unknown) PCR inhibitors by diluting arDNA extracts (see section 3.1.1). If a particular concentration (which might be a tiny amount) of some strong PCR inhibitor was present in the arDNA extract and PCR template, the PCR could be completely blocked regardless of the number of PCR-usable DNA molecules, i.e. amplifiable DNA sequences (see Figure 11; p 117 – PCR experiments with mixing arDNA of *Nezumia* and DNA of control cod DNA). On the other hand, concentrations of DNA and total numbers of PCR-usable DNA molecules/targeted sequences are also important for successful PCR amplifications (Fig. 21(b); p 144).

Theoretical and hypothetical observations might suggest that it is better to use a higher DNA concentration of PCR-DNA templates, i.e. a bigger amount of arDNA in PCR reactions, in order to avoid discrepancies in RAPD-PCR results. This theoretical assumption might be correct according to most of RAPD-PCR results obtained in the study (Fig. 26), but the real situation is much more complex and difficult to bring under the general rule. There are examples from this study that are confusing and/or strongly contradict the assumption about using a bigger volume of PCR-DNA templates as beneficial for RAPD-PCRs (Fig. 27), probably because of a higher concentration of diffusible PCR inhibitors in PCR-DNA

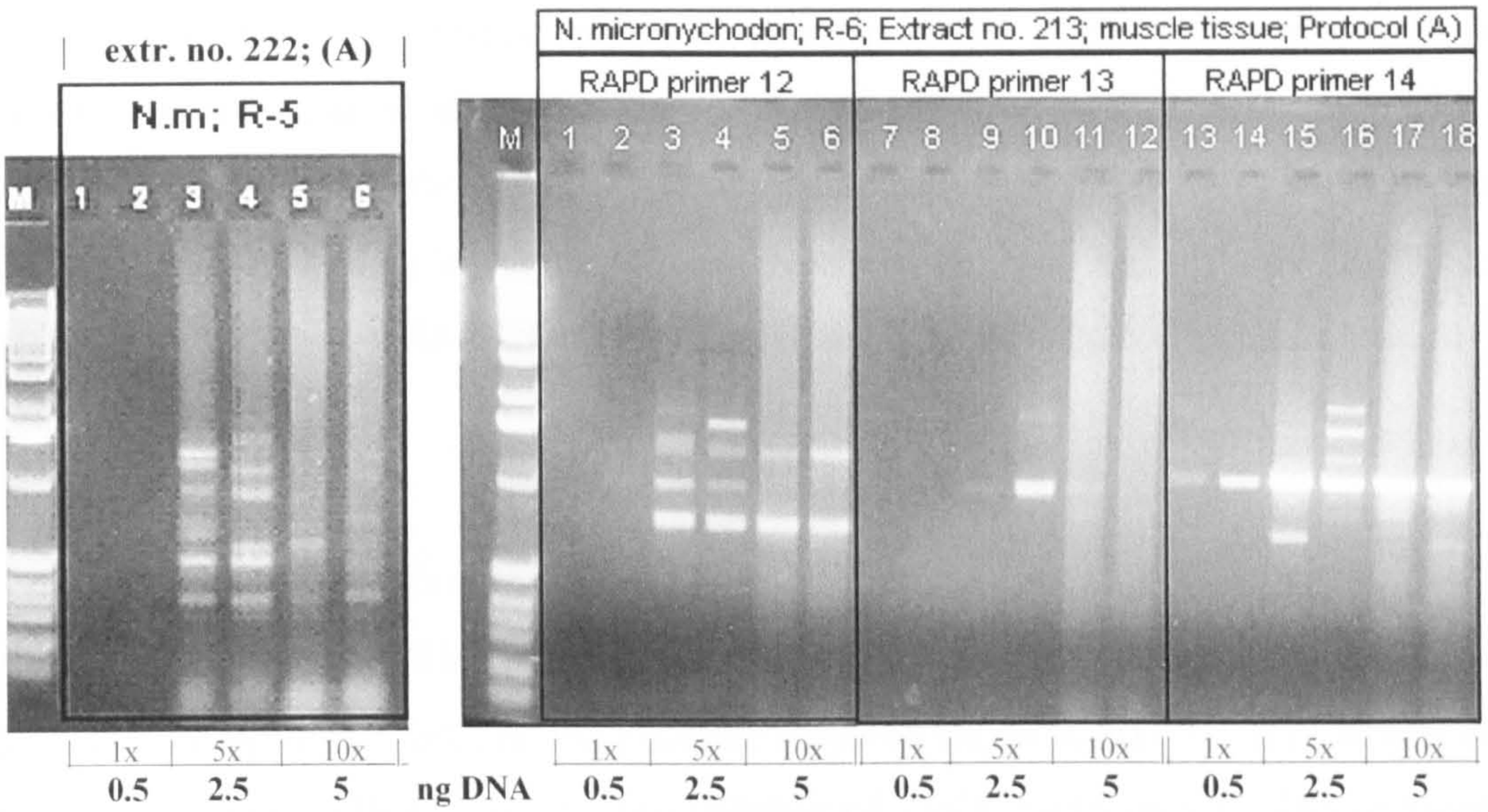
templates (by increasing the volume/amount of arDNA as PCR template, DNA concentration in PCR reaction is increased, but also the concentration of PCR inhibitors).



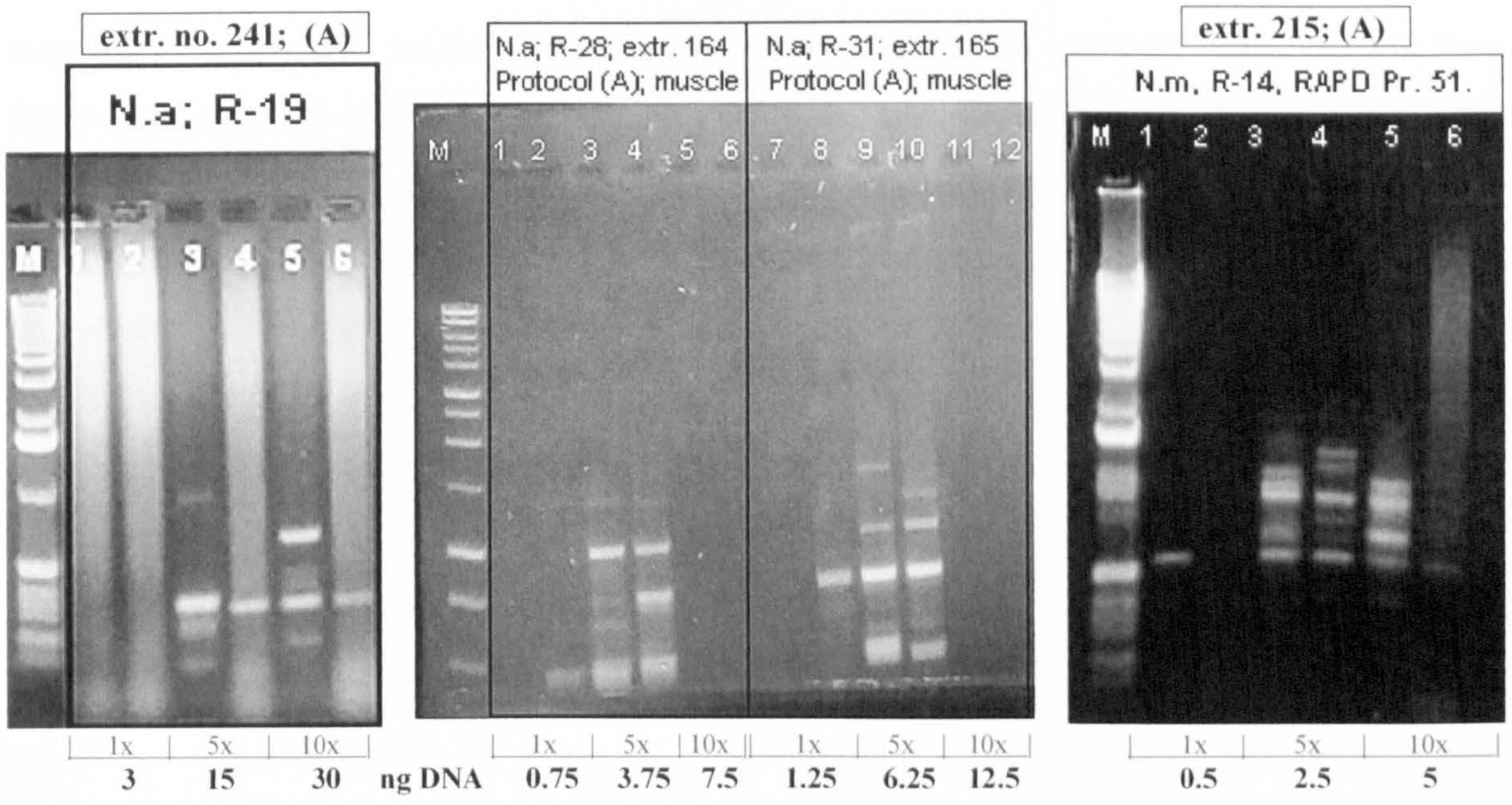

**(a) – RAPD primer 15** **(b) – RAPD primer 37**

**Figure 26** – The increase of PCR efficiency and reproducibility of RAPD-PCR results with increased volume of PCR-DNA templates, that is, concentrations of arDNA in PCR reactions.
"M" indicates size marker fragments - 1 Kb DNA Ladder (Gibco and Promega).
(A) – DNA extracts produced by protocol (A)
R – Identifier of fish individual, i.e. a fish specimen (R-6 is fish no. 6; R-17 is fish no. 17; R-20 is fish no. 20; R-24 is fish no. 24)
N.a – *Nezumia aequalis*
N.m – *Nezumia micronychodon*

Results from this study might suggest that 5-10 ng of arDNA in PCRs will be probably the most suitable amount of arDNA in RAPD-PCRs (Figure 26), but during the study there were many exceptions and confusing results. For example, if 0.5 ng of arDNA in PCR reaction was too little DNA for a successful RAPD-PCR amplification, how could 5 ng be too much DNA for a successful amplification (Fig. 27, gel (a)-Lanes: 1-6, gel (b)-Lanes: 7-12)? It is similar for 7.5 or 12.5 ng DNA in PCR (Fig. 27, gel (d)). These occurrences might be explained by the presence of higher concentrations of diffusible PCR inhibitors in bigger volumes of PCR-DNA templates applied (10x relative concentrations of PCR template: 5 µl of arDNA extract in 25 µl of PCR volume) - compared to 0.5 µl of arDNA extract (1x relative concentration of PCR-DNA template) and 2.5 µl of arDNA extract (5x relative concentration of PCR-DNA template).




**(a) – RAPD primer 15** **(b) – three different RAPD primers**




**(c) – RAPD primer 46** **(d) – RAPD primer 44** **(e) – RAPD primer 51**

**Figure 27** – Confusing and contradictory RAPD-PCR results, that is, increase of PCR efficiency and reproducibility of RAPD-PCR results do not always increase with increased volume of PCR-DNA templates (i.e., concentrations of arDNA in PCR reactions).

It seems that diffusible PCR inhibitors reach some "critical amount" in PCR reactions with which *Taq* polymerase cannot perform amplifications effectively, regardless of the sufficiently available amount of amplifiable DNA, causing a low PCR efficiency with a poor PCR yield, or complete failure of PCR amplification. This might explain the confusing RAPD-PCR results shown in Figure 27 (for example, gels (d) and (e) - the inefficiency, or complete failure, of RAPD-PCRs with 10x PCR-DNA template concentrations in one or both duplicates).

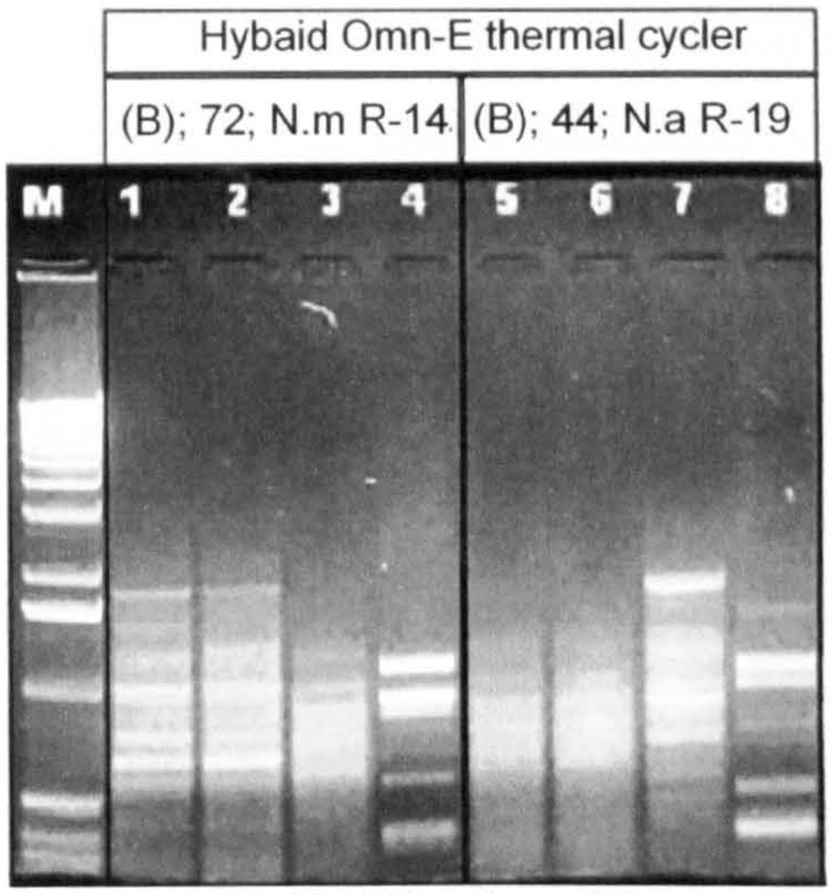
By diluting arDNA extracts with $sdH_2O$ (Figs. 8, 9, 10, 11; pp. 113, 114, 116, 117), the PCR inhibitors also become diluted and it is often possible to achieve a successful PCR amplification despite the fact that the absolute concentrations of arDNA in PCR reactions are lower (by diluting arDNA extracts, the number of DNA molecules is also reduced). The possibility of diluting arDNA extracts and then successfully applying them for RAPD-PCR proves that the main problem is not in DNA concentration and availability of a sufficient number of PCR amplifiable (targeted) sequences in PCRs, i.e. in the damage of arDNA, but in the presence of "diffusible" PCR inhibitors. During the PCR experiments, it was not possible to improve all arDNA extracts just by diluting arDNA extracts, especially not those produced by phenol-based DNA extraction protocols (see section 3.1.1). This emphasises the importance of an adequate selection of DNA extraction protocol for preserved specimens, washing/drying regimes of preserved tissue samples prior to DNA extraction, and the application of DNA purification steps. Other researchers working with difficult samples also found that the PCR inhibition was decreased by diluting DNA extracts until the inhibitor is no longer present at inhibitory concentrations (Boman *et al.* (1999) in clinical samples; Kalmar *et al.* (2000) in ancient bones; Olson *et al.* (2005) – museum dry specimens of skulls and skeletons). However, excessive dilution may also dilute the DNA to non-amplifiable concentrations (Mulligan 2005).

The problems related to the application of the RAPD-PCR methodology to preserved specimens and reproducibility of RAPD-PCR results with formalin/Steedman-arDNA are evident and numerous. Known pitfalls of the RAPD-PCR methodology are much more emphasised with formalin-arDNA than with "good DNA", making this methodology inapplicable for its traditional use on preserved specimens. The expressed difficulties correlate with a particular RAPD primer, arDNA extract, and may be individual fish.

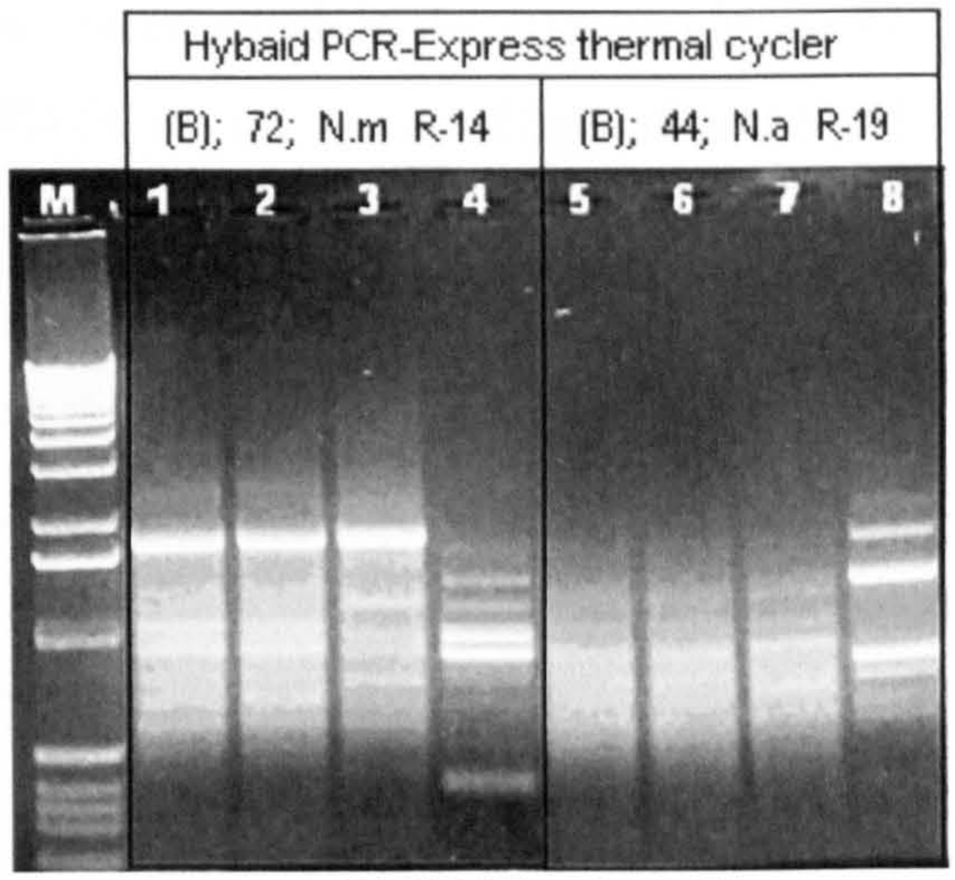
### *3.1.2.2. Taq polymerases*

Before developing the most suitable RAPD-PCR strategy for studying *Nezumia* preserved specimens, the effects of using different *Taq* polymerases, *Taq* buffers, Mg concentrations, additives for improving the PCR performance (DMSO, glycerol, reagents supplied in *Taq* kits as enhancer solutions), PCR thermal cyclers, etc. were tested. The application of a "hot start" (by employing wax beads) and performing RAPD-PCRs in different volumes of PCR reactions (25 µl and 50 µl) were also investigated. It is known that all these factors might (and usually do) cause differences in RAPD-PCR performance and reproducibility of results. These experiments were primarily conducted in order to select the most appropriate RAPD-PCR conditions for investigating formalin-Steedman-arDNA extracts.

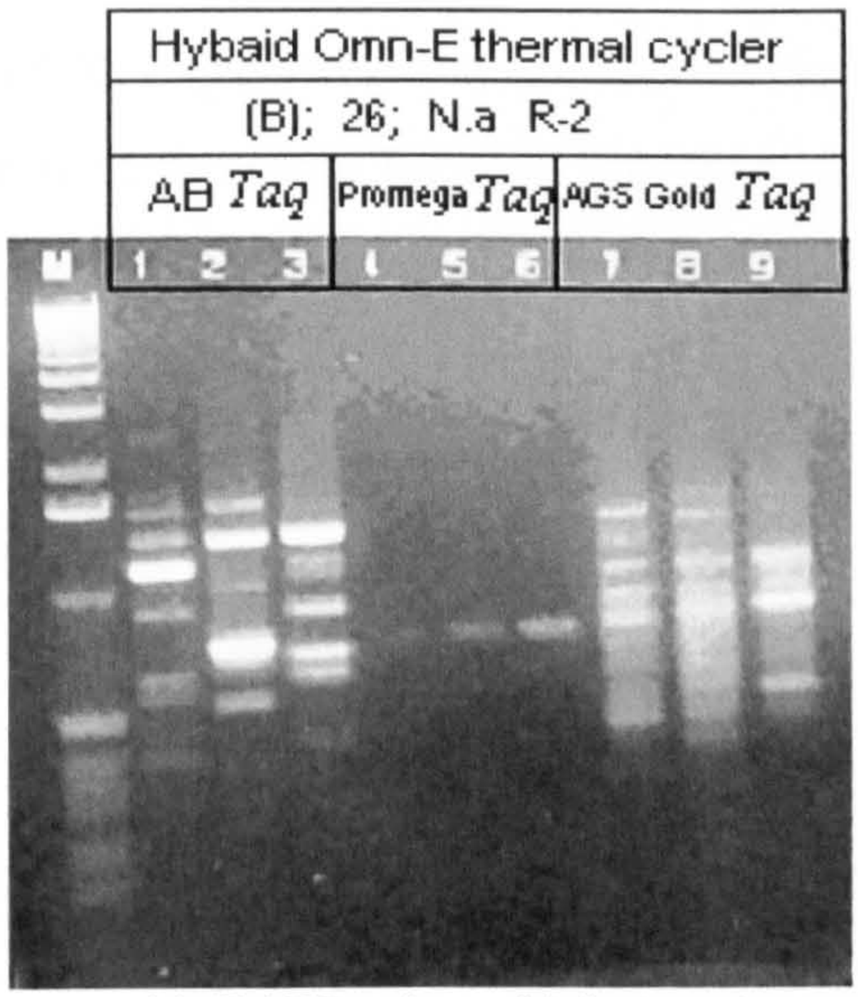
As was expected, the application of different *Taq* polymerases generated different RAPD-PCR profiles (Fig. 28) – compare RAPD-PCR results in gel (a) lanes 1, 2 and 3, RAPD-PCRs which were performed with Hybaid-AGS Gold *Taq* polymerase in supplied Hybaid-AGS Gold Buffer, with the lane 4, RAPD-PCRs performed with Promega *Taq* in Hybaid-AGS Gold Buffer. However, such big differences in RAPD-PCR profiles were not expected with the same arDNA extract. Lanes 1 and 2 are PCR duplicates performed with 2 µl of arDNA suspension [2 ng of DNA] in 25 µl of PCR reaction, whereas RAPD-PCR in lane 3 was performed with 1 µl of arDNA suspension [1 ng DNA] in 25 µl of PCR reaction, i.e. with a half of the DNA amount than in lanes 1 and 2. RAPD-PCR reaction in lane 4 was performed with 2 µl of arDNA suspension [2 ng of DNA], i.e. the PCR reaction had the same content [made from one mastermix] as in lanes 1 and 2, but the only difference was that it was performed with Promega *Taq*). RAPD-PCR experiments presented in lanes 5-8 (Fig. 28) had the same experimental set up as in lanes 1-4, but it was performed with arDNA of different species, i.e. different arDNA extract. At first sight, differences look greater between RAPD-PCRs that were performed with different *Taq* polymerases with the same arDNA extract than with arDNA extracts of different species. The same experiment (the same experimental set up and with the same arDNA extract)) was performed in a different type of thermal cycler (Hybaid PCR-Express with cooling system) – Fig. 28(b). The reproducibility of RAPD-PCRs with different concentrations of arDNA in PCR reactions (with the same *Taq* polymerase) was improved (Lanes: 1-3 and 5-6), but significant differences remained the same between RAPD-PCR results obtained with the same arDNA extracts but performed with different *Taq* polymerases (compare lanes 1, 2 and 3 with lane 4, and lanes 5, 6, 7 with lane 8 in Fig. 28(b)).



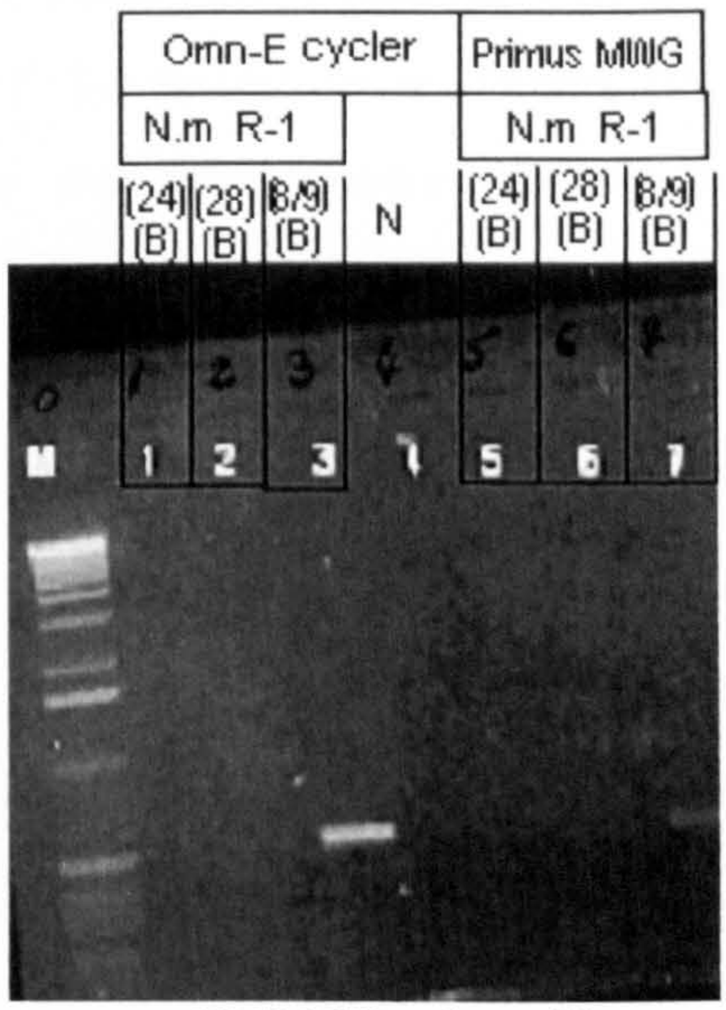

(a) RAPD primer 50




(b) RAPD primer 50




(c) RAPD primer 44; hot start




(d) RAPD primer 29

**Figure 28** – Testing the RAPD-PCR performances and reproducibility of results if different *Taq* polymerases, PCR-DNA template concentrations, and different PCR thermal cyclers were applied.

**Gel (a) and (b)** – Testing different PCR-DNA template concentrations in conjunction with *Taq* polymerases and *Taq* buffers of different suppliers: Hybaid-AGS Gold *Taq* polymerase in supplied Hybaid-AGS Gold *Taq* buffer (Lanes: 1-3, 5-7), and Promega *Taq* polymerase in the AGS Gold *Taq* buffer (Lanes: 4 and 8), and testing different PCR thermal cycler (PCR reactions prepared from the same PCR mastermix were performed in two different thermal cyclers).

**Gel (c)** – Testing different *Taq* polymerases in triplicates (three PCR reactions with the same set up and content of PCR reactions performed in parallel in the same PCR) in conjunction with the "hot start" (employing wax beads)

**Gel (d)** – Testing different PCR thermal cyclers

"M" indicates size marker fragments - 1 Kb DNA Ladder (Gibco), N – negative PCRs,

(B) – DNA extracts produced by protocol (B),

N.a – *Nezumia aequalis*,

N.m – *N. micronychodon*.

From tested *Taq* polymerases, **the Promega *Taq* polymerase** in the supplied Promega *Taq* buffer exhibited the lowest suitability for RAPD-PCRs on the investigated preserved fish specimens because of generating much less RAPD-PCR fragments that were also of very low intensity, i.e. with a poor PCR yield. For example, the Promega *Taq* generated only one band (RAPD fragment of a low PCR yield) with RAPD primer 44 (Fig. 28(c)-L: 4-6). The efficiency and performance of the Promega *Taq* polymerase was significantly improved if Hybaid-AGS Gold Buffer was used instead of supplied Promega *Taq* Buffer (compare the results from gel (c)-Lanes: 4-6 with RAPD-PCR results in gel (a) and (b) – Lanes: 4 and 8). **The Hybaid-AGS Gold *Taq* polymerase** gave reasonably reproducible results (Fig. 28(c)-L: 7-9), but the bands and banding pattern were not clear, i.e. the banding pattern contained smearing regions with a few distinguishable bands which were difficult to count with a high reliability. **The AB Red hot *Taq* polymerase** (Advanced Biotechnologies Ltd., UK) in the supplied AB buffer (Reaction Buffer IV) generated complex band patterns with clear bands (RAPD-PCR fragments), but the reproducibility of the results in triplicates (three PCR reactions with the same content, i.e. made from the same PCR mastermix) was very poor (Fig. 28(c)-L: 1-3). The use of different *Taq* polymerases and *Taq* buffers caused significant differences in RAPD-PCR – in yields of RAPDs and reproducibility and clarity of the RAPD profiles. The sizes of generated RAPDs were affected by the use of different *Taq* polymerases.

The use of different PCR thermocyclers also affected RAPD-PCR results with arDNA. As an example, compare gel (a) and gel (b) in Fig. 28 for differences in the RAPD-PCR results if PCRs were performed in Hybaid Omn-E and Hybaid PCR-Express thermal cyclers (for experimental set up – see description of an experiment in the second paragraph of this section), or compare lanes 1-3 with lanes 5-7 in gel (d) in Fig. 28 for differences in the RAPD-PCR results if PCRs were performed in Hybaid Omn-E and MWG-PCR Primus (MWG-Biotech) thermal cyclers. The use of the Hybaid PCR-Express thermal cycler with cooling system improved the fidelity and specificity of PCR amplifications (Fig. 28, gel (b)) if compared with the results obtained by performing it in Hybaid Omn-E thermal cycler (Fig. 28, gel (a)), whereas the MWG-PCR Primus (MWG-Biotech) did not (Fig 28, gel (d)).

The application of the "hot start" did not significantly improve the RAPD-PCR specificity and reproducibility of PCR products (Fig. 28, gel (c)), but it did improve RAPD-PCR performance related to PCR inhibition (Fig. 29). However, the application of the "hot start" with wax beads was not selected for RAPD-PCR investigations because increasing risk of contamination with external DNA due to additional pipetting and opening tubes in order to add another part of the mastermix after melting, that is, hardening a wax, as well as through handling the wax beads. Instead of applying this risky "hot start", PCR reactions were always placed into a hot thermal cycler block (at 94°C) in order to increase the specificity of PCRs. The application of polymerase-antibody enzymes ("antibody hot start") is worthwhile to test and apply in research with arDNAs.

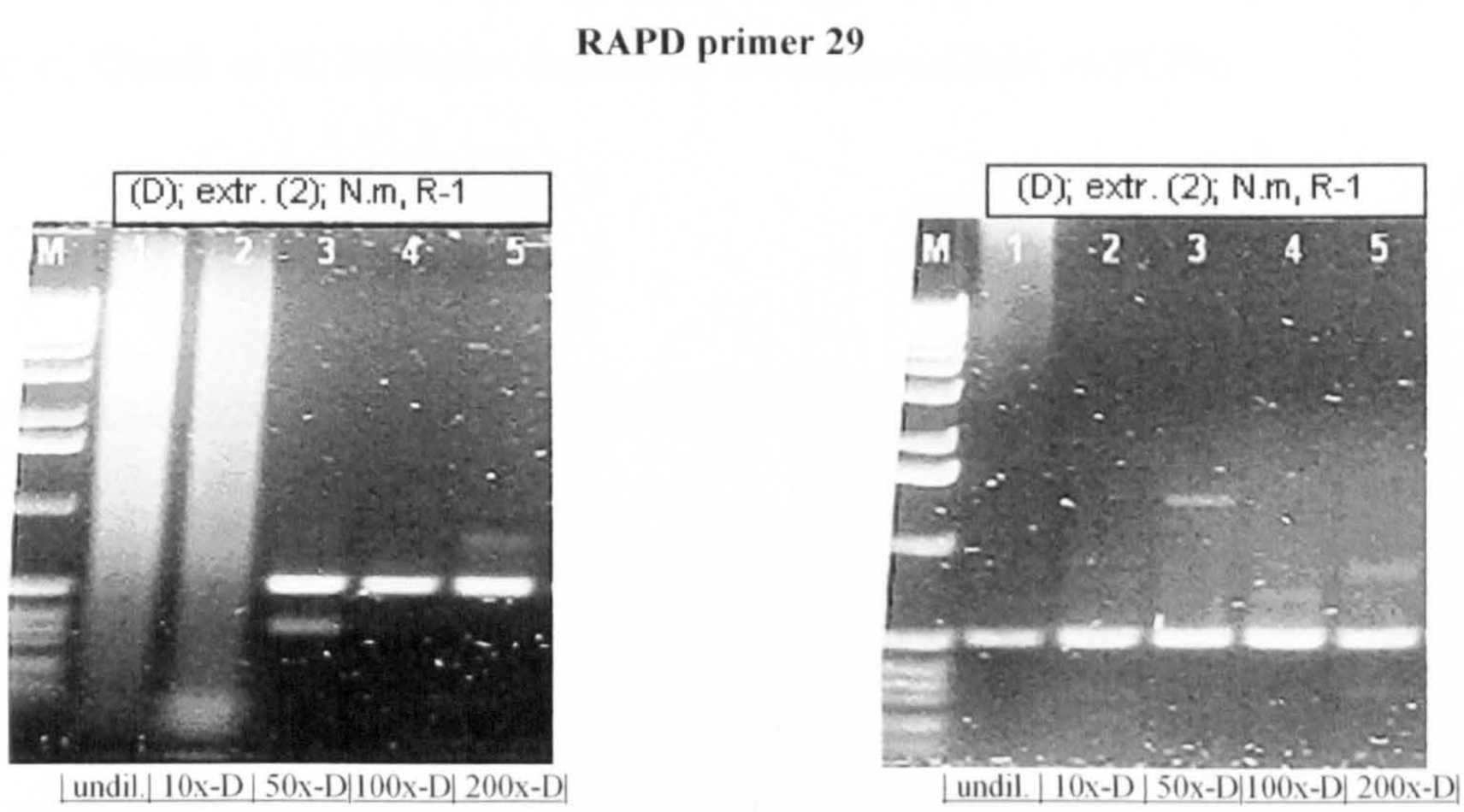



**Figure 29** – ***Testing the effect of "hot start" on the PCR inhibition.*** The arDNA extracts produced by modified protocol (D) exhibited the PCR inhibition as undiluted arDNA suspension (gel (a)-Lane: 1) and 10x-diluted suspension (gel (a)-Lane: 2) if "hot start" was not applied. However, if "hot start" was applied (employing wax beads), the RAPD product (band size of ~500 bp) was generated (gel (b)-Lanes: 1 and 2). If 50x, 100x and 200x dilutions of the arDNA extract (50x-D, 100x-D and 200x-D) were applied in PCR reactions, the RAPD product was generated regardless of applying "hot start" or not (compare gels (a) and (b) - Lanes: 3, 4, 5). In all PCRs, Promega *Taq* was applied and 2.5 µl of arDNA suspension (diluted or undiluted) in 50 µl of PCR reaction volume. All PCR reactions were run simultaneously in the Hybaid Omn-E thermal cycler.
"M" indicates size marker fragments - 1 Kb DNA Ladder (Gibco).
(D) – arDNA extracts produced by protocol (D),
N.m – *N. micronychodon.*
R – Identifier of fish individual, i.e. a fish specimen (R-1 is fish no. 1)

After extensive investigations and optimisations of RAPD-PCRs with different RAPD primers and arDNA extracts, the Hybaid Omn-E thermal cycler and Haybaid-AGS Gold *Taq* (with supplied AGS Gold *Taq* Buffer - a buffer without Mg) proved to be the most suitable for RAPD-PCR experiments in this study. They were therefore selected as standards for RAPD-PCRs in this project. In order to ensure the reliability and accuracy of obtained RAPD-PCR results, it was found that the most reliable RAPD-PCR results were obtained if each RAPD-PCR experiment was performed with three different PCR-DNA template concentrations in duplicates of each template concentration.

The data suggest that the appropriate choice of *Taq* polymerase is important for the PCR amplification success rate with arDNA, PCR amplification yield and reproducibility of PCR products, as well as the size of PCR products. Similar findings were also reported by other researchers (e.g., Quach *et al.* 2004) for the use of formalin-arDNA in PCRs.

### *3.1.2.3. RAPD experiments on preserved specimens of Nezumia*

One intention was to investigate the possibility of using the RAPD-PCR methodology (different RAPD primers) on degraded/damaged DNA (formalin/Steedman's-arDNA of *Nezumia aequalis* and *N. micronychodon*) by using developed RAPD strategy, i.e. performing RAPD-PCRs with three different PCR-DNA template concentrations in duplicates. The aim was to investigate if any of tested RAPD primers could be applicable to formalin/Steedman's-arDNA of these two *Nezumia* species for a traditional RAPD-distance analysis, as well as to find a possible presence of species-specific band(s) and/or species-specific RAPD profiles (band patterns). In this study, 4,428 RAPD-PCRs, applying 60 different primers, were performed (Table 7).

**Table 7.** – RAPD-PCRs with formalin-arDNA of *Nezumia aequalis* (N.a) and *N. micronychodon* (N.m)

| RAPD Primer | Species | $n_1$ | $n_2$ | Total no. of PCRs | Success of RAPD-PCRs | | | | | | Number of bands | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | "+" amplif. | % | "+/–" amplif. | % | "–" amplif. | % | Range | Average |
| Primer 1. | N.m | 3 | 3 | 13 | 3 | 23 | 0 | 0 | 10 | 77 | 2 - 5 | 3.50 |
| | N.a | 3 | 3 | 13 | 10 | 77 | 0 | 0 | 3 | 23 | 1 - 9 | 6.10 |
| | **N.m + N.a** | **6** | **6** | **26** | **13** | **50** | **0** | **0** | **13** | **50** | **1 - 9** | |
| Primer 2. | N.m | 6 | 8 | 48 | 19 | 40 | 0 | 0 | 29 | 60 | 1 - 10 | 3.10 |
| | N.a | 13 | 13 | 85 | 43 | 51 | 2 | 2 | 40 | 47 | 1 - 10 | 3.95 |
| | **N.m + N.a** | **19** | **21** | **133** | **62** | **47** | **2** | **2** | **69** | **51** | **1 - 10** | |
| Primer 3. | N.m | 3 | 3 | 13 | 8 | 62 | 0 | 0 | 5 | 38 | 5 - 8 | 6.75 |
| | N.a | 3 | 3 | 13 | 12 | 92 | 0 | 0 | 1 | 8 | 8 - 12 | 10.08 |
| | **N.m + N.a** | **6** | **6** | **26** | **20** | **77** | **0** | **0** | **6** | **23** | **5 - 12** | |
| Primer 4. | N.m | 2 | 2 | 12 | 11 | 92 | 0 | 0 | 1 | 8 | 1 - 8 | 4.50 |
| | N.a | 3 | 3 | 13 | 13 | 100 | 0 | 0 | 0 | 0 | 1 -12 | 6.38 |
| | **N.m + N.a** | **5** | **5** | **25** | **24** | **96** | **0** | **0** | **1** | **4** | **1 - 12** | |
| Primer 5. | N.m | 3 | 3 | 13 | 6 | 46 | 0 | 0 | 7 | 54 | 1 - 5 | 2.50 |
| | N.a | 3 | 3 | 13 | 11 | 85 | 0 | 0 | 2 | 15 | 2 - 8 | 5.55 |
| | **N.m + N.a** | **6** | **6** | **26** | **17** | **65** | **0** | **0** | **9** | **35** | **1 - 8** | |
| Primer 6. | N.m | 3 | 3 | 13 | 9 | 69 | 0 | 0 | 4 | 31 | 1 - 6 | 3.33 |
| | N.a | 3 | 3 | 13 | 12 | 92 | 0 | 0 | 1 | 8 | 1 - 10 | 6.29 |
| | **N.m + N.a** | **6** | **6** | **26** | **21** | **81** | **0** | **0** | **5** | **19** | **1 - 10** | |
| Primer 7. | N.m | 3 | 3 | 13 | 10 | 77 | 1 | 8 | 2 | 15 | 1 - 10 | 5.90 |
| | N.a | 3 | 3 | 13 | 13 | 100 | 0 | 0 | 0 | 0 | 1 - 9 | 5.54 |
| | **N.m + N.a** | **6** | **6** | **26** | **23** | **88** | **1** | **4** | **2** | **8** | **1 - 10** | |
| Primer 8. | N.m | 3 | 3 | 13 | 2 | 15 | 1 | 8 | 10 | 77 | 1 - 2 | 1.40 |
| | N.a | 3 | 3 | 13 | 12 | 92 | 0 | 0 | 1 | 8 | 2 - 7 | 4.54 |
| | **N.m + N.a** | **6** | **6** | **26** | **14** | **54** | **1** | **4** | **11** | **42** | **1 - 7** | |
| Primer 9. | N.m | 3 | 3 | 13 | 7 | 54 | 0 | 0 | 6 | 46 | 1 - 6 | 2.14 |
| | N.a | 3 | 3 | 13 | 12 | 92 | 0 | 0 | 1 | 8 | 2 - 8 | 5.58 |
| | **N.m + N.a** | **6** | **6** | **26** | **19** | **73** | **0** | **0** | **7** | **27** | **1 - 8** | |
| Primer 10. | N.m | 3 | 3 | 14 | 1 | 7 | 0 | 0 | 13 | 93 | 1 - 1 | 1.00 |
| | N.a | 3 | 3 | 14 | 12 | 86 | 0 | 0 | 2 | 14 | 1 - 13 | 5.25 |
| | **N.m + N.a** | **6** | **6** | **28** | **13** | **46** | **0** | **0** | **15** | **54** | **1 - 13** | |
| Primer 11. | N.m | 11 | 15 | 86 | 30 | 35 | 0 | 0 | 56 | 65 | 1 - 8 | 3.50 |
| | N.a | 17 | 17 | 110 | 69 | 63 | 0 | 0 | 41 | 37 | 1 - 11 | 5.00 |
| | **N.m + N.a** | **28** | **32** | **196** | **99** | **51** | **0** | **0** | **97** | **49** | **1 - 11** | |
| Primer 12. | N.m | 3 | 3 | 14 | 11 | 79 | 0 | 0 | 3 | 21 | 1 - 5 | 2.91 |
| | N.a | 3 | 3 | 14 | 12 | 86 | 0 | 0 | 2 | 14 | 4 - 10 | 5.92 |
| | **N.m + N.a** | **6** | **6** | **28** | **23** | **82** | **0** | **0** | **5** | **18** | **1 - 10** | |
| Primer 13. | N.m | 3 | 3 | 13 | 4 | 31 | 1 | 8 | 8 | 61 | 1 - 2 | 1.50 |
| | N.a | 3 | 3 | 13 | 7 | 54 | 0 | 0 | 6 | 46 | 1 - 3 | 1.42 |
| | **N.m + N.a** | **6** | **6** | **26** | **11** | **42** | **1** | **4** | **14** | **54** | **1 - 3** | |
| Primer 14. | N.m | 3 | 3 | 13 | 11 | 85 | 0 | 0 | 2 | 15 | 2 - 5 | 3.00 |
| | N.a | 3 | 3 | 13 | 12 | 92 | 0 | 0 | 1 | 8 | 3 - 7 | 5.08 |
| | **N.m + N.a** | **6** | **6** | **26** | **23** | **88** | **0** | **0** | **3** | **12** | **2 - 7** | |
| Primer 15. | N.m | 11 | 19 | 98 | 64 | 65 | 0 | 0 | 34 | 35 | 1 - 12 | 3.95 |
| | N.a | 17 | 20 | 116 | 103 | 89 | 0 | 0 | 13 | 11 | 1 - 9 | 4.89 |
| | **N.m + N.a** | **28** | **39** | **214** | **167** | **78** | **0** | **0** | **47** | **22** | **1 - 12** | |

Table 7 – continued

| RAPD Primer | Species | $n_1$ | $n_2$ | Total no. of PCRs | Success of RAPD-PCRs | | | | | | Number of bands | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | "+" amplif. | % | "+/–" amplif. | % | "–" amplif. | % | Range | Average |
| Primer 16. | N.m | 2 | 2 | 7 | 3 | 43 | 0 | 0 | 4 | 57 | 4 - 6 | 5.00 |
| | N.a | 3 | 3 | 13 | 11 | 85 | 0 | 0 | 2 | 15 | 1 - 8 | 5.21 |
| | **N.m + N.a** | **5** | **5** | **20** | **14** | **70** | **0** | **0** | **6** | **30** | **1 - 8** | |
| Primer 17. | N.m | 2 | 2 | 7 | 2 | 29 | 0 | 0 | 5 | 71 | 5 - 5 | 5.00 |
| | N.a | 3 | 3 | 13 | 10 | 77 | 0 | 0 | 3 | 23 | 1 - 6 | 4.10 |
| | **N.m + N.a** | **5** | **5** | **20** | **12** | **60** | **0** | **0** | **8** | **40** | **1 - 6** | |
| Primer 18. | N.m | 2 | 2 | 7 | 4 | 57 | 1 | 14 | 2 | 29 | 3 - 8 | 5.00 |
| | N.a | 3 | 3 | 13 | 11 | 85 | 2 | 15 | 0 | 0 | 1 - 8 | 5.63 |
| | **N.m + N.a** | **5** | **5** | **20** | **15** | **75** | **3** | **15** | **2** | **10** | **1 - 8** | |
| Primer 19. | N.m | 2 | 2 | 7 | 3 | 43 | 1 | 14 | 3 | 43 | 2 - 2 | 2.00 |
| | N.a | 3 | 3 | 13 | 12 | 92 | 0 | 0 | 1 | 8 | 2 - 10 | 6.17 |
| | **N.m + N.a** | **5** | **5** | **20** | **15** | **75** | **1** | **5** | **4** | **20** | **2 - 10** | |
| Primer 20. | N.m | 2 | 2 | 7 | 5 | 71 | 0 | 0 | 2 | 29 | 2 - 6 | 3.60 |
| | N.a | 3 | 3 | 13 | 12 | 92 | 0 | 0 | 1 | 8 | 1 - 16 | 9.00 |
| | **N.m + N.a** | **5** | **5** | **20** | **17** | **85** | **0** | **0** | **3** | **15** | **1 - 16** | |
| Primer 21. | N.m | 3 | 3 | 13 | 12 | 92 | 1 | 8 | 0 | 0 | 2 - 10 | 6.75 |
| | N.a | 3 | 3 | 13 | 13 | 100 | 0 | 0 | 0 | 0 | 5 - 11 | 7.62 |
| | **N.m + N.a** | **6** | **6** | **26** | **25** | **96** | **1** | **4** | **0** | **0** | **2 - 11** | |
| Primer 22. | N.m | 3 | 9 | 24 | 9 | 38 | 1 | 4 | 14 | 58 | 1 - 9 | 4.44 |
| | N.a | 3 | 3 | 13 | 12 | 92 | 0 | 0 | 1 | 8 | 2 - 12 | 6.92 |
| | **N.m + N.a** | **6** | **12** | **37** | **21** | **57** | **1** | **3** | **15** | **40** | **1 - 12** | |
| Primer 23. | N.m | 3 | 3 | 13 | 1 | 8 | 0 | 0 | 12 | 92 | 9 - 9 | 9.00 |
| | N.a | 3 | 3 | 13 | 2 | 15 | 0 | 0 | 11 | 85 | 1 - 1 | 1.00 |
| | **N.m + N.a** | **6** | **6** | **26** | **3** | **12** | **0** | **0** | **23** | **88** | **1 - 9** | |
| Primer 24. | N.m | 3 | 3 | 13 | 0 | 0 | 4 | 31 | 9 | 69 | ? 1 | ? |
| | N.a | 3 | 3 | 13 | 10 | 77 | 0 | 0 | 3 | 23 | 1 - 10 | 3.30 |
| | **N.m + N.a** | **6** | **6** | **26** | **10** | **38** | **4** | **15** | **12** | **46** | **1 - 10** | |
| Primer 25. | N.m | 3 | 3 | 13 | 0 | 0 | 1 | 8 | 12 | 92 | ? 2 | ? |
| | N.a | 3 | 3 | 13 | 4 | 31 | 3 | 23 | 6 | 46 | 1 - 4 | 3.40 |
| | **N.m + N.a** | **6** | **6** | **26** | **4** | **15** | **4** | **15** | **18** | **70** | **1 - 4** | |
| Primer 26. | N.m | 3 | 3 | 13 | 0 | 0 | 0 | 0 | 13 | 100 | 0 | 0 |
| | N.a | 3 | 3 | 13 | 10 | 77 | 0 | 0 | 3 | 23 | 1 - 7 | 4.50 |
| | **N.m + N.a** | **6** | **6** | **26** | **10** | **38** | **0** | **0** | **16** | **62** | **1 - 7** | |
| Primer 27. | N.m | 3 | 3 | 13 | 0 | 0 | 1 | 8 | 12 | 92 | ? 1 | ? |
| | N a | 3 | 3 | 13 | 6 | 46 | 0 | 0 | 7 | 54 | 1 - 1 | 1.00 |
| | **N.m + N.a** | **6** | **6** | **26** | **6** | **23** | **1** | **4** | **19** | **73** | **1 - 1** | |
| Primer 28. | N.m | 4 | 5 | 19 | 13 | 68 | 0 | 0 | 6 | 32 | 1 - 10 | 6.00 |
| | N.a | 3 | 3 | 13 | 13 | 100 | 0 | 0 | 0 | 0 | 4 - 10 | 6.77 |
| | **N.m + N.a** | **7** | **8** | **32** | **26** | **81** | **0** | **0** | **6** | **19** | **1 - 10** | |
| Primer 29. | N.m | 12 | 87 | 293 | 102 | 35 | 8 | 3 | 183 | 62 | 1 - 13 | 3.41 |
| | N.a | 19 | 66 | 181 | 34 | 19 | 2 | 1 | 145 | 80 | 1 - 10 | 4.54 |
| | **N.m + N.a** | **31** | **153** | **474** | **136** | **29** | **10** | **2** | **328** | **69** | **1 - 13** | |
| Primer 30. | N.m | 3 | 3 | 13 | 11 | 85 | 0 | 0 | 2 | 15 | 1 - 7 | 4.55 |
| | N a | 3 | 3 | 13 | 12 | 92 | 0 | 0 | 1 | 8 | 2 - 7 | 4.33 |
| | **N.m + N.a** | **6** | **6** | **26** | **23** | **88** | **0** | **0** | **3** | **12** | **1 - 7** | |
| Primer 31. | N.m | 3 | 3 | 13 | 12 | 92 | 0 | 0 | 1 | 8 | 2 - 14 | 6.92 |
| | N.a | 3 | 3 | 13 | 13 | 100 | 0 | 0 | 0 | 0 | 1 - 12 | 9.00 |
| | **N.m + N.a** | **6** | **6** | **26** | **25** | **96** | **0** | **0** | **1** | **4** | **1 - 14** | |
| Primer 32. | N.m | 6 | 14 | 51 | 46 | 90 | 1 | 2 | 4 | 8 | 1 - 9 | 5.37 |
| | N.a | 7 | 16 | 51 | 49 | 96 | 0 | 0 | 2 | 4 | 1 - 12 | 5.92 |
| | **N.m + N.a** | **13** | **30** | **102** | **95** | **93** | **1** | **1** | **6** | **6** | **1 - 12** | |

**Table 7 – continued**

| RAPD Primer | Species | $n_1$ | $n_2$ | Total no. of PCRs | Success of RAPD-PCRs | | | | | | Number of bands | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | "+" amplif. | % | "+/–" amplif. | % | "–" amplif. | % | Range | Average |
| Primer 33. | N.m | 3 | 3 | 13 | 4 | 31 | 0 | 0 | 9 | 69 | 3 - 7 | 5.25 |
| | N.a | 3 | 3 | 13 | 12 | 92 | 0 | 0 | 1 | 8 | 1 - 9 | 5.00 |
| | **N.m + N.a** | **6** | **6** | **26** | **16** | **62** | **0** | **0** | **10** | **38** | **1 - 9** | |
| Primer 34. | N.m | 3 | 3 | 13 | 11 | 85 | 0 | 0 | 2 | 15 | 1 - 9 | 4.09 |
| | N.a | 3 | 3 | 13 | 12 | 92 | 0 | 0 | 1 | 8 | 4 - 11 | 7.17 |
| | **N.m + N.a** | **6** | **6** | **26** | **23** | **88** | **0** | **0** | **3** | **12** | **1 - 11** | |
| Primer 35. | N.m | 3 | 3 | 13 | 9 | 69 | 0 | 0 | 4 | 31 | 2 - 11 | 6.67 |
| | N.a | 3 | 3 | 13 | 12 | 92 | 0 | 0 | 1 | 8 | 2 - 7 | 5.33 |
| | **N.m + N.a** | **6** | **6** | **26** | **21** | **81** | **0** | **0** | **5** | **19** | **2 - 11** | |
| Primer 37. | N.m | 11 | 16 | 97 | 36 | 37 | 0 | 0 | 61 | 63 | 1 - 11 | 5.08 |
| | N.a | 17 | 17 | 109 | 76 | 70 | 0 | 0 | 33 | 30 | 1 - 11 | 5.64 |
| | **N.m + N.a** | **28** | **33** | **206** | **112** | **54** | **0** | **0** | **94** | **46** | **1 - 11** | |
| Primer 38. | N.m | 3 | 3 | 13 | 12 | 92 | 0 | 0 | 1 | 8 | 1 - 15 | 7.67 |
| | N.a | 3 | 3 | 13 | 13 | 100 | 0 | 0 | 0 | 0 | 1 - 13 | 4.85 |
| | **N.m + N.a** | **6** | **6** | **26** | **25** | **96** | **0** | **0** | **1** | **4** | **1 - 15** | |
| Primer 39. | N.m | 3 | 3 | 13 | 10 | 77 | 0 | 0 | 3 | 23 | 1 - 10 | 3.80 |
| | N.a | 3 | 3 | 13 | 12 | 92 | 0 | 0 | 1 | 8 | 1 - 8 | 5.17 |
| | **N.m + N.a** | **6** | **6** | **26** | **22** | **85** | **0** | **0** | **4** | **15** | **1 - 10** | |
| Primer 40. | N.m | 2 | 2 | 12 | 3 | 25 | 0 | 0 | 9 | 75 | 1 - 4 | 2.34 |
| | N.a | 2 | 2 | 12 | 11 | 92 | 0 | 0 | 1 | 8 | 1 - 10 | 3.82 |
| | **N.m + N.a** | **4** | **4** | **24** | **14** | **58** | **0** | **0** | **10** | **42** | **1 - 10** | |
| Primer 41. | N.m | 1 | 3 | 5 | 4 | 80 | 1 | 20 | 0 | 0 | 1 - 4 | 2.50 |
| | N.a | 1 | 3 | 3 | 2 | 67 | 0 | 0 | 1 | 33 | 2 - 5 | 3.50 |
| | **N.m + N.a** | **2** | **6** | **8** | **6** | **75** | **1** | **12.5** | **1** | **12.5** | **1 - 5** | |
| Primer 42. | N.m | 8 | 9 | 44 | 23 | 52 | 0 | 0 | 21 | 48 | 1 - 9 | 4.26 |
| | N.a | 8 | 10 | 45 | 37 | 82 | 0 | 0 | 8 | 18 | 1 - 10 | 6.24 |
| | **N.m + N.a** | **16** | **19** | **89** | **60** | **67** | **0** | **0** | **29** | **33** | **1 - 10** | |
| Primer 43. | N.m | 1 | 3 | 5 | 0 | 0 | 0 | 0 | 5 | 100 | 0 | 0 |
| | N.a | 1 | 3 | 3 | 0 | 0 | 0 | 0 | 3 | 100 | 0 | 0 |
| | **N.m + N.a** | **2** | **6** | **8** | **0** | **0** | **0** | **0** | **8** | **100** | **0** | **0** |
| Primer 44. | N.m | 12 | 72 | 466 | 245 | 52 | 3 | 1 | 218 | 47 | 1 - 11 | 3.59 |
| | N a | 19 | 99 | 611 | 373 | 61 | 11 | 2 | 227 | 37 | 1 - 12 | 4.19 |
| | **N.m + N.a** | **31** | **171** | **1077** | **618** | **58** | **14** | **1** | **445** | **41** | **1 - 12** | |
| Primer 45. | N.m | 5 | 8 | 30 | 21 | 70 | 0 | 0 | 9 | 30 | 1 - 11 | 5.00 |
| | N.a | 5 | 7 | 27 | 24 | 89 | 0 | 0 | 3 | 11 | 1 - 12 | 5.79 |
| | **N.m + N.a** | **10** | **15** | **57** | **45** | **79** | **0** | **0** | **12** | **21** | **1 - 12** | |
| Primer 46. | N.m | 11 | 17 | 135 | 72 | 53 | 3 | 2 | 60 | 45 | 1 - 8 | 2.93 |
| | N.a | 16 | 19 | 166 | 133 | 80 | 1 | 1 | 32 | 19 | 1 - 12 | 4.07 |
| | **N.m + N.a** | **27** | **36** | **301** | **205** | **68** | **4** | **1** | **92** | **31** | **1 - 12** | |
| Primer 47. | N.m | 4 | 6 | 20 | 16 | 80 | 0 | 0 | 4 | 20 | 1 - 9 | 3.13 |
| | N.a | 4 | 7 | 20 | 16 | 80 | 0 | 0 | 4 | 20 | 1 - 13 | 6.25 |
| | **N.m + N.a** | **8** | **13** | **40** | **32** | **80** | **0** | **0** | **8** | **20** | **1 - 13** | |
| Primer 48. | N.m | 11 | 26 | 101 | 43 | 43 | 1 | 1 | 57 | 56 | 1 - 10 | 3.46 |
| | N.a | 15 | 30 | 138 | 78 | 56 | 1 | 1 | 59 | 43 | 1 - 14 | 5.73 |
| | **N.m + N.a** | **26** | **56** | **239** | **121** | **51** | **2** | **1** | **116** | **48** | **1 - 14** | |

Table 7 – continued

| RAPD Primer | Species | $n_1$ | $n_2$ | Total no. of PCRs | Success of RAPD-PCRs | | | | | | Number of bands | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | "+" amplif. | % | "+/–" amplif. | % | "–" amplif. | % | Range | Average |
| Primer 49. | N.m | 4 | 18 | 46 | 36 | 78 | 0 | 0 | 10 | 22 | 1 - 9 | 4.44 |
| | N.a | 4 | 16 | 44 | 42 | 95 | 0 | 0 | 2 | 5 | 1 - 9 | 5.30 |
| | **N.m + N.a** | **8** | **34** | **90** | **78** | **87** | **0** | **0** | **12** | **13** | **1 - 9** | |
| Primer 50. | N.m | 3 | 5 | 22 | 18 | 82 | 0 | 0 | 4 | 18 | 1 - 10 | 6.67 |
| | N.a | 3 | 4 | 21 | 19 | 90 | 0 | 0 | 2 | 10 | 3 - 11 | 6.32 |
| | **N.m + N.a** | **6** | **9** | **43** | **37** | **86** | **0** | **0** | **6** | **14** | **1 - 11** | |
| Primer 51. | N.m | 4 | 5 | 17 | 15 | 88 | 1 | 6 | 1 | 6 | 1 - 9 | 4.93 |
| | N a | 3 | 4 | 16 | 15 | 94 | 1 | 6 | 0 | 0 | 2 - 9 | 5.00 |
| | **N.m + N.a** | **7** | **9** | **33** | **30** | **91** | **2** | **6** | **1** | **3** | **1 - 9** | |
| Primer 52. | N.m | 6 | 15 | 28 | 12 | 43 | 1 | 4 | 15 | 53 | 1 - 9 | 4.67 |
| | N.a | 6 | 13 | 26 | 20 | 77 | 0 | 0 | 6 | 23 | 2 - 13 | 9.84 |
| | **N.m + N.a** | **12** | **28** | **54** | **32** | **59** | **1** | **2** | **21** | **39** | **1 - 13** | |
| Primer 53. | N.m | 3 | 7 | 19 | 14 | 74 | 0 | 0 | 5 | 26 | 1 - 8 | 3.79 |
| | N.a | 5 | 5 | 17 | 16 | 94 | 0 | 0 | 1 | 6 | 1 - 13 | 7.38 |
| | **N.m + N.a** | **8** | **12** | **36** | **30** | **83** | **0** | **0** | **6** | **17** | **1 - 13** | |
| Primer 54. | N.m | 3 | 3 | 13 | 5 | 38 | 0 | 0 | 8 | 62 | 1 - 4 | 2.40 |
| | N.a | 3 | 3 | 13 | 12 | 92 | 0 | 0 | 1 | 8 | 6 - 13 | 9.17 |
| | **N.m + N.a** | **6** | **6** | **26** | **17** | **65** | **0** | **0** | **9** | **35** | **1 - 13** | |
| Primer 55. | N.m | 3 | 3 | 13 | 10 | 77 | 0 | 0 | 3 | 23 | 1 - 9 | 5.50 |
| | N.a | 3 | 3 | 13 | 13 | 100 | 0 | 0 | 0 | 0 | 3 - 11 | 8.15 |
| | **N.m + N.a** | **6** | **6** | **26** | **23** | **88** | **0** | **0** | **3** | **12** | **1 - 11** | |
| Primer 56. | N.m | 3 | 3 | 13 | 8 | 62 | 0 | 0 | 5 | 38 | 1 - 7 | 3.75 |
| | N.a | 3 | 3 | 13 | 13 | 100 | 0 | 0 | 0 | 0 | 5 - 12 | 8.23 |
| | **N.m + N.a** | **6** | **6** | **26** | **21** | **81** | **0** | **0** | **5** | **19** | **1 - 12** | |
| Primer 57. | N.m | 3 | 3 | 13 | 6 | 46 | 0 | 0 | 7 | 54 | 1 - 4 | 2.60 |
| | N.a | 3 | 3 | 13 | 11 | 85 | 0 | 0 | 2 | 15 | 2 - 12 | 8.18 |
| | **N.m + N.a** | **6** | **6** | **26** | **17** | **65** | **0** | **0** | **9** | **35** | **1 - 12** | |
| Primer 58. | N.m | 3 | 3 | 13 | 8 | 62 | 0 | 0 | 5 | 38 | 1 - 10 | 6.25 |
| | N.a | 3 | 3 | 13 | 12 | 92 | 1 | 8 | 0 | 0 | 2 - 11 | 8.83 |
| | **N.m + N.a** | **6** | **6** | **26** | **20** | **77** | **1** | **4** | **5** | **19** | **1 - 11** | |
| Primer 59. | N.m | 3 | 3 | 13 | 4 | 31 | 2 | 15 | 7 | 54 | 1 - 2 | 1.75 |
| | N.a | 3 | 3 | 13 | 7 | 54 | 1 | 8 | 5 | 38 | 1 - 5 | 3.14 |
| | **N.m + N.a** | **6** | **6** | **26** | **11** | **42** | **3** | **12** | **12** | **46** | **1 - 5** | |
| Primer 60. | N.m | 3 | 3 | 13 | 11 | 85 | 0 | 0 | 2 | 15 | 1 - 11 | 5.73 |
| | N.a | 3 | 3 | 13 | 12 | 92 | 0 | 0 | 1 | 8 | 2 - 11 | 7.25 |
| | **N.m + N.a** | **6** | **6** | **26** | **23** | **88** | **0** | **0** | **3** | **12** | **1 - 11** | |
| All RAPD-PCRs (Primers 1-60) TOTAL | N.m | | | 2118 | 1075 | 51 | 34 | 2 | 1009 | 47 | 1 - 15 | 4.08 |
| | N.a | | | 2310 | 1590 | 69 | 25 | 2 | 695 | 29 | 1 - 16 | 5.59 |
| | **N.m + N.a** | | | **4428** | **2665** | **60** | **59** | **2** | **1704** | **38** | **1 - 16** | |

N.m - *Nezumia micronychodon*

N.a - *Nezumia aequalis*

$n_1$ - number of fish individuals tested with the particular RAPD primer

$n_2$ - number of formalin-arDNA extracts tested with the particular RAPD primer

**"+" amplif.** – successful PCR amplifications, i.e. generated RAPD-PCR products (at least one clear RAPD-PCR fragment)

**"+/–" amplif.** – a questionable success of PCR amplifications (generated bands are faint and may be artifacts)

**"–" amplif.** – unsuccessful PCR amplifications, i.e. RAPD-PCR products were not generated

With respect to the executed tests (considering both investigated species), 60% RAPD-PCRs successfully generated PCR products, 2% were of a questionable success (generated unclear, very faint band/s which might not be genuine PCR products, i.e. it is a high possibility that those bands were artifactual), and 38% were unsuccessful RAPD-PCRs (PCR products were not generated). A successful RAPD-PCR amplification was considered if at least one clear band (RAPD-PCR fragment) was generated. The formalin-arDNA of *Nezumia aequalis* gave a higher percentage of successful RAPD-PCRs (69%) than formalin-arDNA of *N. micronychodon* (51%). Also, on average, a higher number of RAPD-PCR fragments (bands) were generated with formalin-Steedman's-arDNA extracts of *Nezumia aequalis* (5.59) than of *N. micronychodon* (4.08). These may indicate better conditions and PCR suitability of *Nezumia aequalis* formalin-arDNAs than that of *N. micronychodon*, but the differences also might be caused by differences in genome sequences of the two investigated species.

The estimation of the success rate with a particular RAPD primer might help in investigating the genetic structure of species with unstudied genomes and provide information on the usefulness of particular RAPD primers to be used on formalin-preserved specimens and in further research with particular species (two *Nezumia* species in this study). For example, RAPD primer 26 generated PCR products with formalin/Steedman-arDNAs of *N. aequalis*, but not with formalin/Steedman's-arDNAs of *N. micronychodon*. A similar situation was with the application of RAPD primers 24, 25, 27. Some RAPD primers did not generate any PCR product with formalin-arDNAs of both species (e.g., primer 43), or the tested primers gave very poor amplifications with non-reproducible bands (e.g. primer 23). This kind of information could be useful for drawing certain conclusions about *N. micronychodon* and *N. aequalis* genomes. If some RAPD primers did not produce successful PCR amplifications, the reason might not be in the damage of DNA due to preservation, but rather a different genomic structure of investigated species; that is, that these primers did not have a good match in genuine DNA sequences in one of the two, or both, investigated species.

The results with RAPD primer 43 need to be viewed with caution because of a low number of PCRs, although it was tested with three different arDNA extracts in both species. However, the findings with primers 24 and 26 are most probably genuine and indicate molecular differences between these two species. These two RAPD primers gave successful RAPD-PCR amplifications with almost all tested arDNA extracts of *N. aequalis*, but none with arDNA of *N. micronychodon*. This ratio of successful and unsuccessful RAPD-PCR

amplifications with RAPD primer 24 (none for *N. micronychodon* and 77% of successful PCR amplifications for *N. aequalis*) suggests genuine differences in the particular DNA genome regions of the two species. Moreover, a possible further confirmation of the validity of these findings is the fact that two arDNA extracts of *N. micronychodon* (extracts: 213 and 215) did not give RAPD-PCR products with mentioned RAPD primers, but produced successful amplifications with other RAPD-PCR primers (e.g. with primers 28 and 29) in the same PCR run and PCR reactions that were made from the same PCR master mix (see Figs. 21(b) and 22; pp 144 and 145). Results are probably representative and indicate differences in DNA sequences of the two investigated species, but the possibility cannot be excluded that particular arDNA extracts used in RAPD-PCR experiments gave false negative PCRs with arDNA of *N. micronychodon* because of particular damaged regions of DNA sequences, or for other reasons (see detailed explanations of the "untypical and non-standardised behaviour" of different formalin-arDNA extracts in the previous sections).

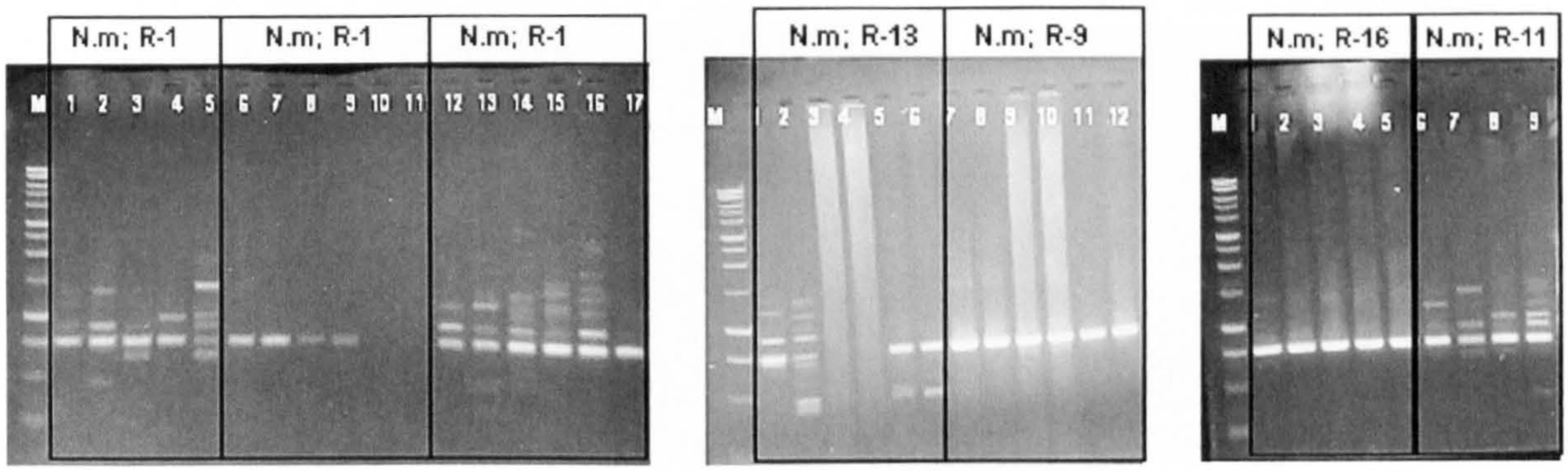
A higher rate of successful RAPD-PCRs in particular species (for instance primers 8 and 22 for *N. aequalis*, or primers 4 and 32 for both species) might indicate a potentially higher usefulness of these RAPD primers to be used on preserved specimens (because of the possibility that a higher amplification success rate indicates genome regions that are less susceptible to DNA damage caused by preservation and other factors, and as a result are more appropriate for PCR amplifications with a particular RAPD primer).

With regard to the reproducibility and the consistency of RAPDs, RAPDs were generally inconsistent, but some of RAPD-PCR fragments were reproducible with particular RAPD primers (e.g. with RAPD primer 44 bands sizes of ~750 bp and 500 bp; or with primer 46 bands sizes ~700 bp and ~300 bp - Figs. 30 and 31) with the majority of tested arDNA extracts. In general, some RAPDs were reproducible with all tested PCR-DNA template concentrations, some RAPD-PCR fragments were reproducible only with a particular DNA template concentration, and some were reproducible only in duplicates with the same DNA template concentration (regardless of the concentration of PCR-DNA templates applied) - see Figs. 30 and 31 and the support material on CD. However, some bands (RAPD-PCR fragments) were completely unreliable – not only were they inconsistent and irreproducible across different PCR experiments and different DNA-PCR template concentrations, but they were not reproducible even in their duplicates. This kind of RAPD result (sometimes reproducible, sometimes not reproducible) occurred frequently in the study, making comparison and interpretation of results difficult. Besides the inconsistency of RAPD-PCR

results in general (related to the presence of RAPD-PCR fragments in profiles), the intensity of co-migrating bands was often different, even in PCR duplicates. This clearly indicates that RAPDs are erratic and definitely cannot be used with formalin-Steedman's archival DNA for their traditional applications (presence/absence of bands and RAPD-distance analysis). Developed strategy (performing RAPD-PCRs with three different PCR-DNA template concentrations in duplicates) aided better understanding of the problems related to the use of formalin-Steedman's-arDNA in RAPD-PCRs and prevented misinterpretation of results, but could not lead to its reliable traditional application (i.e. reliable scoring of bands for accurate RAPD-distance analysis) on investigated specimens.

However, the RAPD-PCR methodology can be useful in investigating genetic structure and isolating markers from formalin-fixed specimens and species with unstudied genomes (in this study from the two deep-sea fish - *Nezumia aequalis* and *N. micronychodon*). This is based on results from this study which clearly demonstrated that some of obtained RAPD-PCR fragments (bands) were unquestionably genuine. The example that a robust RAPD primer could produce consistent RAPD-PCR fragments with formalin-fixed specimens has already been mentioned - RAPD primer 44 with the RAPD-DNA fragment size of ~750 bp (Fig. 30) and RAPD primer 46 with fragment size of ~700 (Fig. 31). The RAPD-PCR fragments sizes of 350 bp and 1200 bp generated with RAPD primer 46 were also reproducible and shared in both species, but these two bands were not consistently reproducible across all experiments (Fig. 31). The degree of their reproducibility was often in correlation with a particular archival DNA extract and PCR-DNA template concentration used in PCR experiments. It is important to emphasise that without performing Southern blot and/or sequencing experiments, shared, co-migrating bands do not necessarily mean that they belong to the same locus (region) in the genomes. For *Nezumia* species, the homology and identity of a co-migrating and shared band size of ~750 bp (generated with primer 44) was confirmed by Southern blotting (Fig. 32) and by sequencing (see section 3.2). In order to investigate the specificity of this RAPD marker generated for *Nezumia aequalis* and *N. micronychodon*, the RAPD-PCR profiles of these two investigated species were compared with RAPD-PCR profiles of closely related species (*Coryphaenoides armatus*; Fig. 17(a); p 126) and distant fish species (rainbow trout; Figures 13 and 17; pp 121 and 126). The RAPD profiles were distinctly different, although the band size of 750-800 bp was also present in rainbow trout (Fig. 13; p 121).

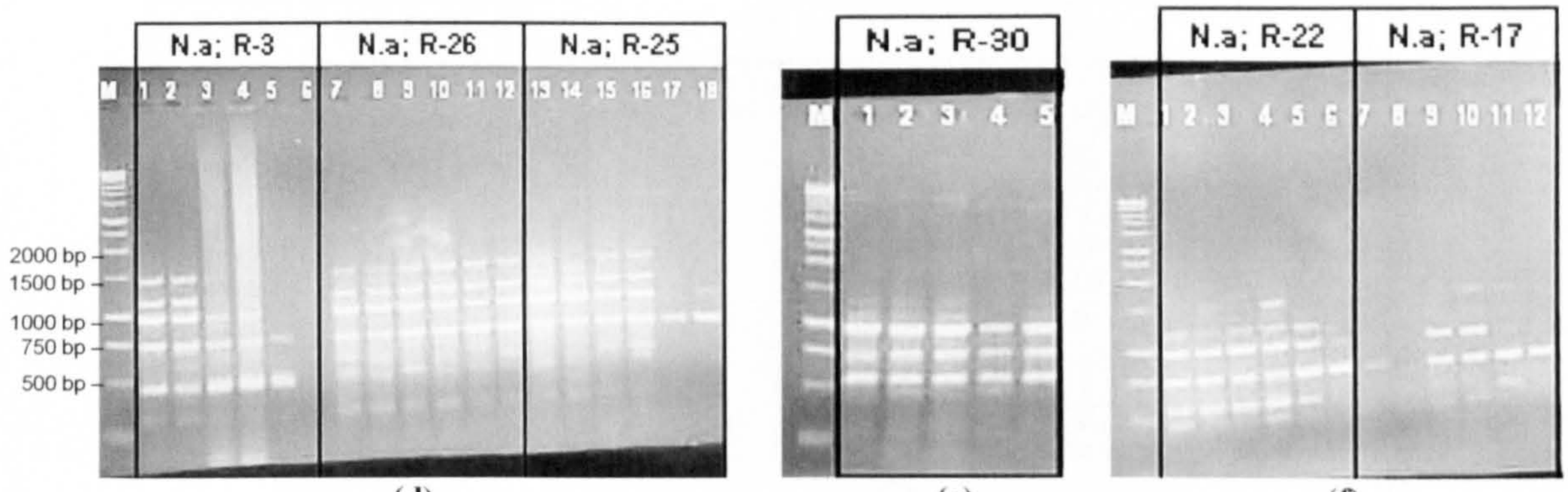
*1) Nezumia micronychodon*



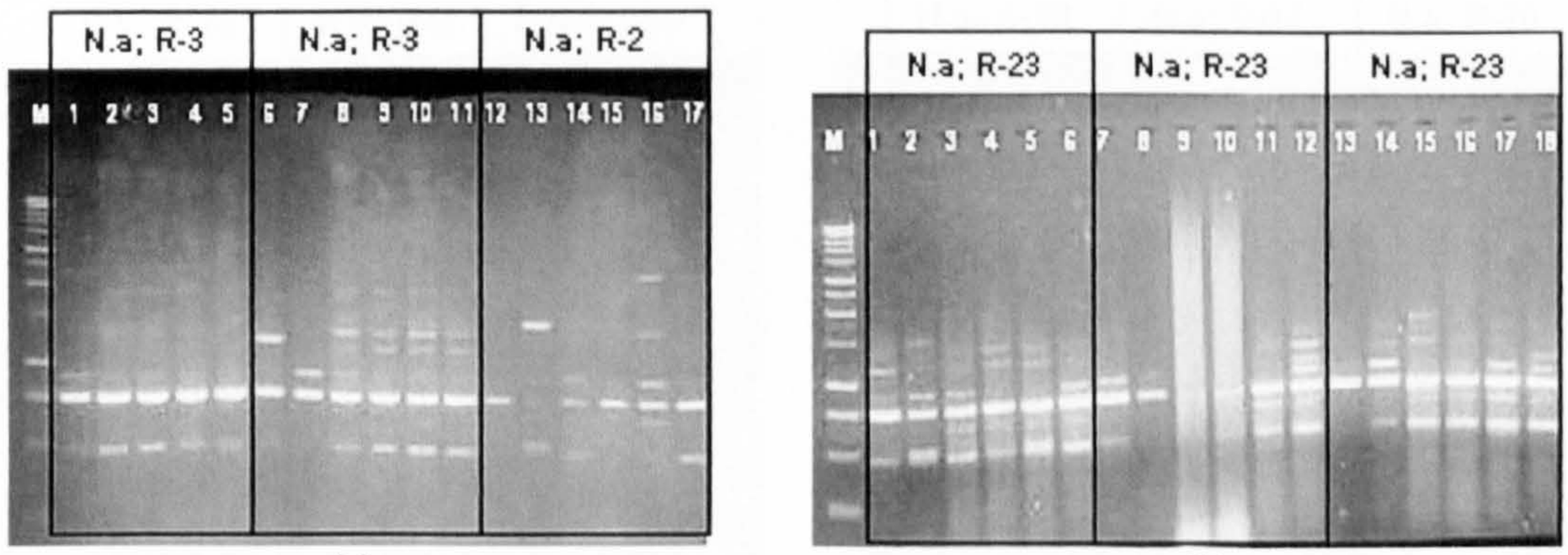

**(a)** **(b)** **(c)**

---

*2) Nezumia aequalis*



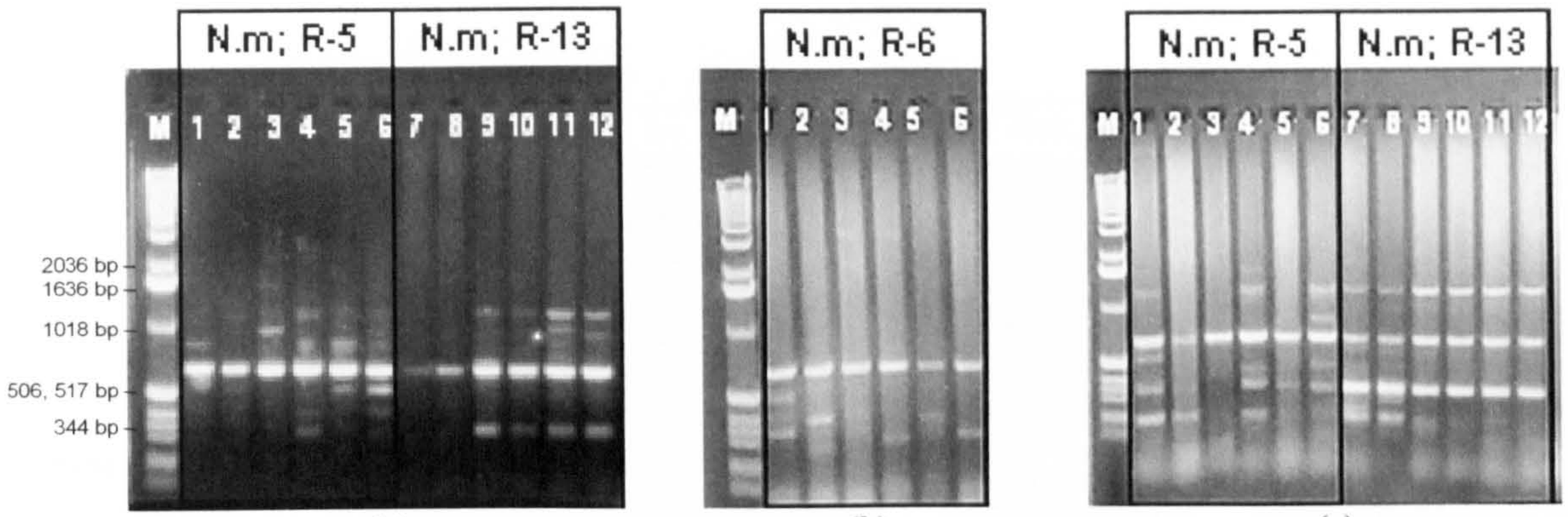

**(d)** **(e)** **(f)**




**(g)** **(h)**

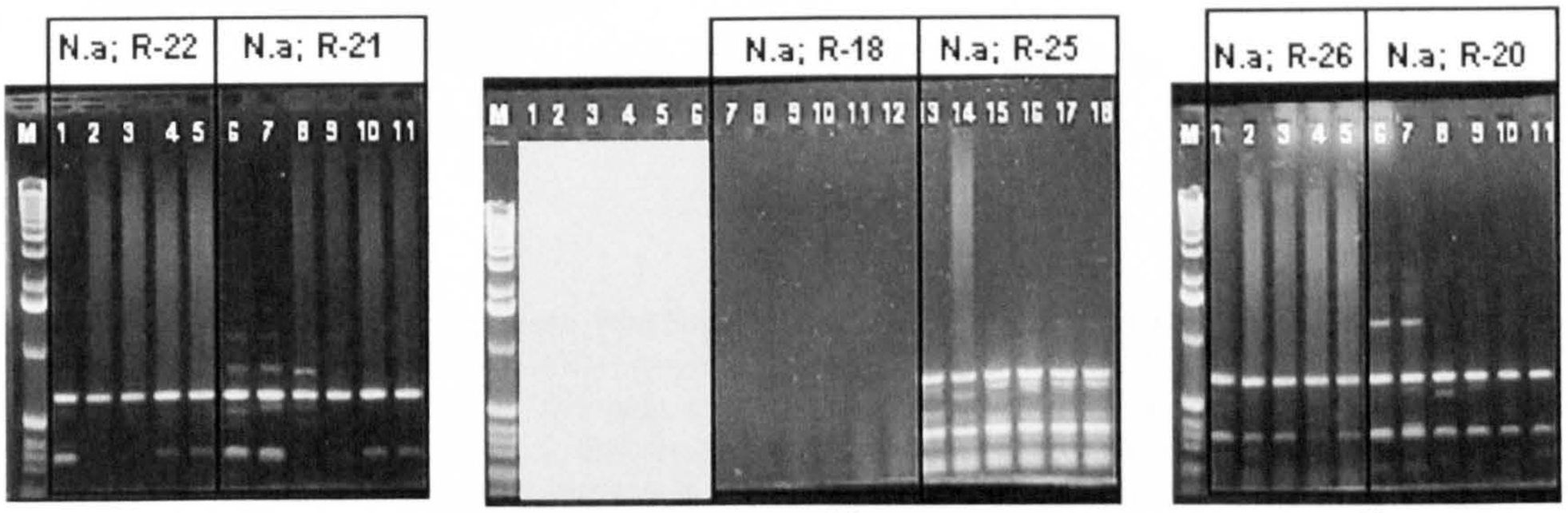
**Figure 30** – RAPD-PCR experiments with formalin-arDNAs of *N. micronychodon* (N.m) and *N. aequalis* (N.a) performed with RAPD Primer 44, different arDNA extracts from different fish individuals. The band size of ~750 bp was reproducible throughout experiments regardless of fish individuals, species and arDNA extracts used in these RAPD-PCR experiments. This 750-bp RAPD fragment is shared band in both species. Southern blotting (Fig. 32) proved the homology of this RAPD between the two investigated species.
"M" indicates size marker fragments - 1 Kb DNA Ladder (Promega).
R – Identifier of fish individual, i.e. a fish specimen (e.g., R-1 is fish no. 1; R-3 is fish no. 3, etc.)

*1) Nezumia micronychodon*




---

*2) Nezumia aequalis*



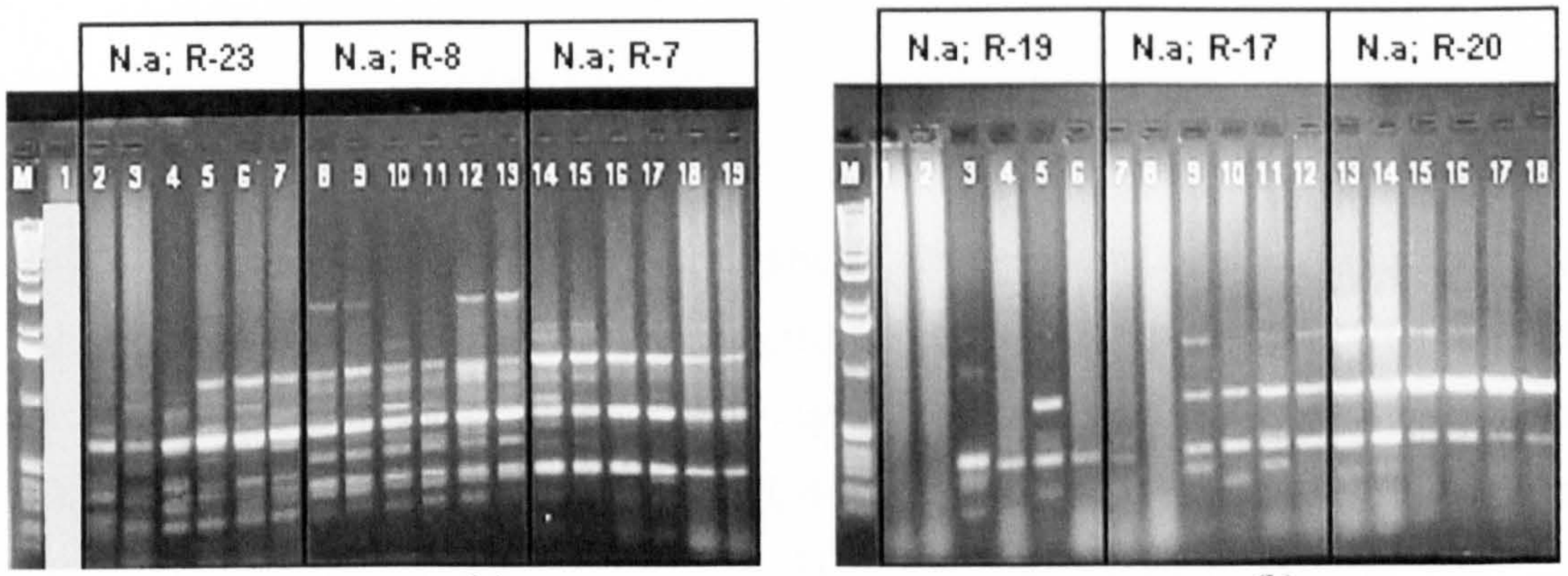




**Figure 31** – RAPD-PCR experiments with formalin-arDNAs of *N. micronychodon* (N.m) and *N. aequalis* (N.a) performed with RAPD Primer 46, different arDNA extracts from different fish individuals. The band size of ~700 bp was reproducible throughout experiments regardless of fish individuals, species and arDNA extracts used in these RAPD-PCR experiments. This 700-bp RAPD fragment is shared band in both species. The bands sizes of ~350 bp and ~1200 bp are also reproducible, but not as consistently as a 700 bp-band. These two RAPDs are present in both species as shared RAPD-fragments.
"M" indicates size marker fragments - 1 Kb DNA Ladder (Gibco).
R – Identifier of fish individual, i.e. fish specimen (e.g., R-5 is fish no. 5; R-23 is fish no. 23)

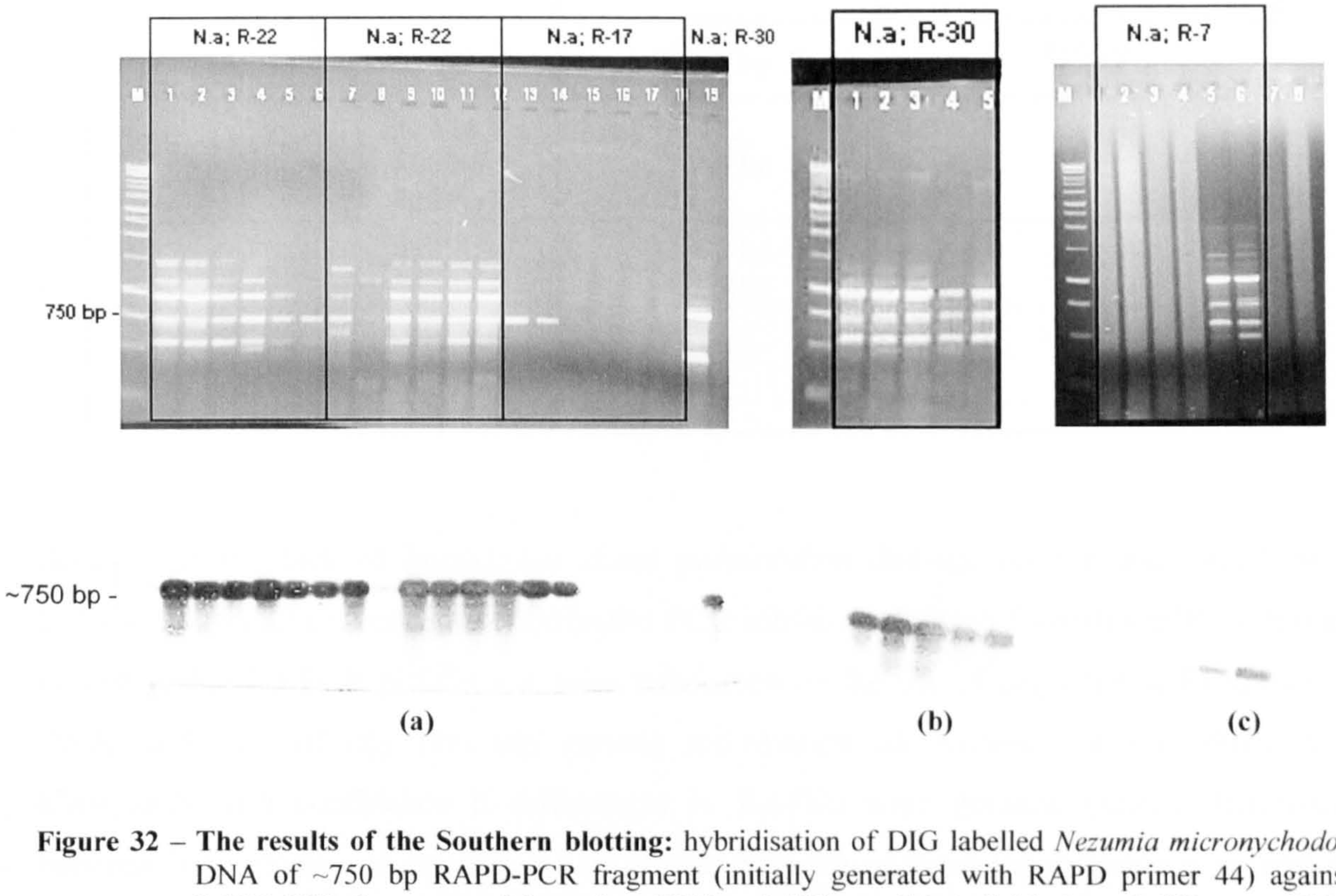



**Figure 32 – The results of the Southern blotting:** hybridisation of DIG labelled *Nezumia micronychodon* DNA of ~750 bp RAPD-PCR fragment (initially generated with RAPD primer 44) against RAPD-PCR fragments of formalin-arDNAs of different fish individuals and different arDNA extracts of *N. aequalis* (N.a). Southern blot results indicated strong match with all tested RAPD-PCR fragments of the same size produced by RAPD primer 44. This confirmed homology of the comigrating RAPD fragments size of 750 bp across the experiments and species.
"M" indicates size marker fragments - 1 Kb DNA Ladder (Promega).

Besides RAPD primers 44 and 46, there were other RAPD primers that gave informative and reasonably reproducible results with the majority of tested arDNA extracts (for example, primers 11 and 15 - see Table 7 and 8, and support material on the CD). Some possible *Nezumia aequalis*-specific and *N. micronychodon*-specific RAPD fragments, as well as some shared (co-migrating) bands, with particular RAPD primers are shown in Table 8.

**Table 8** – Possible species-specific and shared bands with particular RAPD primers

| RAPD primer | Shared bands | *N. aequalis*-specific | *N. micronychodon*-specific |
|---|---|---|---|
| **11** | | 1000 bp | ~850 bp |
| **15** | 350 bp<br>800 bp<br>900 bp | 1500 bp | |
| **44** | ~750 bp | 1000 bp<br>1500 bp | |
| **46** | ~350 bp<br>~700 bp<br>~1200 bp | | |

Because of the lack of knowledge about preservation damage on formalin-Steedman's-arDNA and the exact content of diffusible PCR inhibitors in each formalin-arDNA extract, as well as RAPD-PCR pitfalls that were reinforced by the use of degraded and fragmented DNA, and lack of any previous genetic information on *Nezumia*, it was difficult to distinguish with confidence if differences in RAPDs were genuine genetic differences between individuals and species, or if it was because of artificial differences caused by DNA preservation and its different recovery in different arDNA extracts. There is no published data on estimation of artifactual RAPD fragments from preserved specimens. This was the reason that in this study there were no attempts to carry out the conventional RAPD analysis by applying standard algorithms and statistical calculations. This was to prevent possible misinterpretation of results until the mentioned issues related to preserved specimens are better understood. Sequencing of particular RAPD-PCR fragments and converting RAPDs into STS and SCAR markers is the only reliable approach with archival DNA (see section 3.2).

### 3.1.3. Possible relationships in molecular work with archival DNA

Based on empirical data from this study, there are many possible relationships to consider in work with formalin-Steedman-arDNA. For instance:

- DNA extraction protocol/ PCR amplifiable arDNA/ PCR success rate/ fish individual,
- DNA extraction protocol/ washing and drying a tissue sample prior to DNA extraction/ the presence of PCR inhibitors,

- DNA extraction protocol/ PCR marker system,
- DNA extraction protocol/ the success with a particular RAPD primer/ PCR-DNA template concentration,
- DNA extraction protocol/ RAPD primer – reproducibility of results;
- DNA extraction protocol/ possibility to dilute arDNA extracts in order to make arDNA usable for the PCRs
- DNA extraction protocol/ reproducibility of a particular RAPD-PCR fragments (bands)/ reproducibility of RAPD profiles/ number of bands in RAPD-PCR profiles,
- size of bands/ number of bands/ PCR-DNA template volume (i.e. relative DNA concentrations of templates)/ diluted arDNA extracts, etc.

Most of these relationships are not "independent", i.e. they are related to each other (directly and indirectly), although a DNA extraction protocol is the main connection in this net of relationships (Fig. 33).

The strongest link is between DNA extraction protocol and the ability to produce PCR amplifiable arDNA. There is also some kind of relationship between DNA extraction protocol and pre-extraction treatment, on the one hand, and their applicability for particular PCR marker system, on the other. Clear examples of these relationships are shown in Figs. 15 and 19 (pp 124 and 130). These relationships are not completely understood, but PCR experiments from this study strongly suggest some kind of connection.

Absolute concentration of arDNA in PCR reaction is important for the success of PCR amplification, but it seems that the volume of arDNA suspension in PCRs (i.e. relative concentration of PCR-DNA templates) is much more important (probably because of the presence of diffusible PCR inhibitors). Good arDNA extracts produce RAPD-PCR products with a whole range of absolute concentrations of arDNA in PCRs (from 0.08 ng to 40 ng; Fig. 18; p 128, or even 90 ng – Fig. 24(e)-L: 4-6; p 147) and relative PCR-DNA template concentrations (1x, 5x and 10x, i.e. 0.5 µl, 2.5 µl and 5 µl of arDNA suspension in 25 µl of PCR reaction). However, for the success of mitochondrial PCR amplifications, it seems that a selection of a DNA extraction protocol and an absolute concentration of arDNA in PCR reactions are both very important. Is it because of the presence of highly damaged mitochondrial arDNA, or because of a very small proportion of mtDNA in extracts of a total genomic arDNA, or for some another reason? These questions are still without definite answers (see section 3.3).

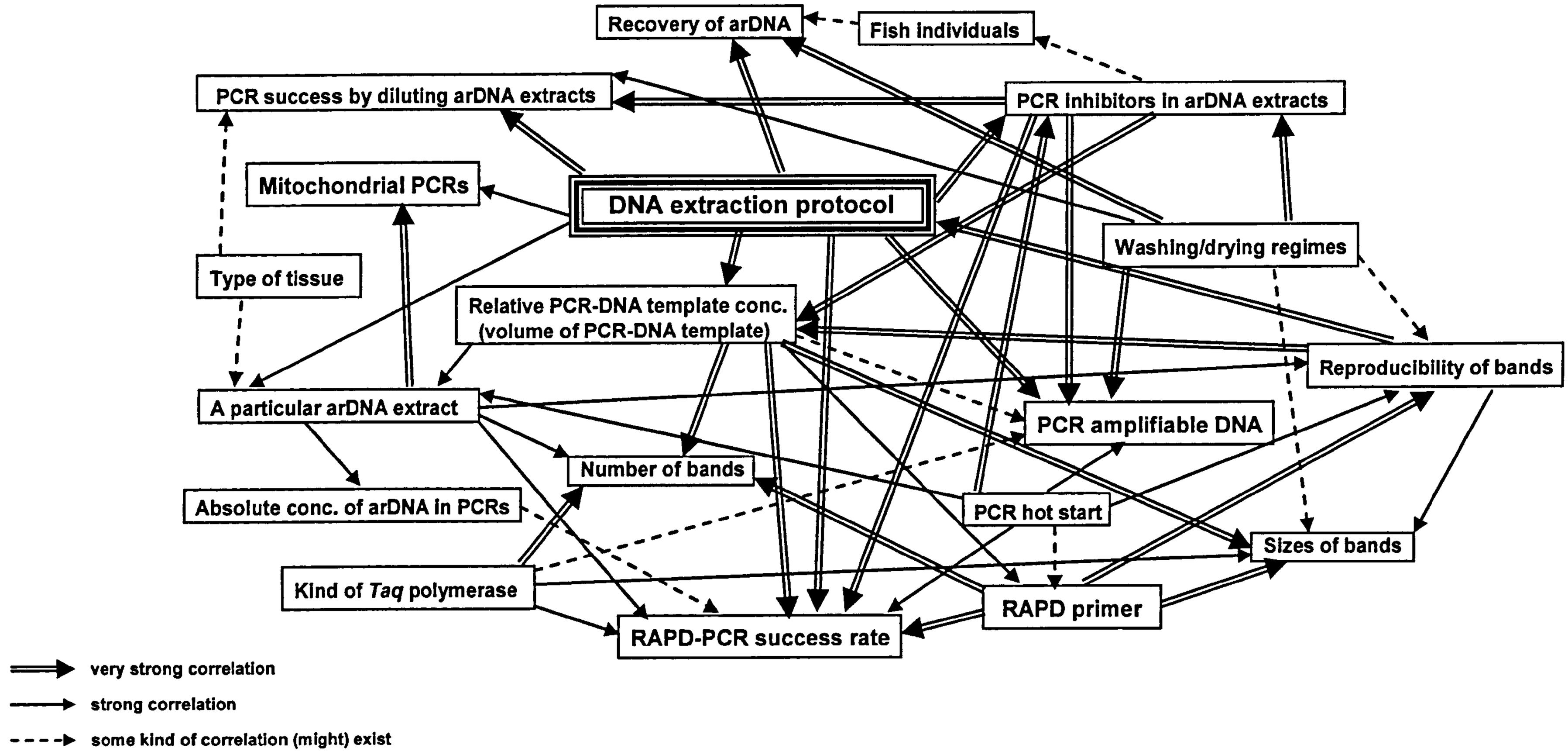



**Figure 33** – "Relationships net" based on empirical data – direct and indirect relationships

Crucial points for successful molecular work with already existing preserved specimens are as follows (in order as significance):

1. DNA extraction protocol
2. (a) Pre-extraction treatments of preserved tissue (washing/drying a tissue sample)
   (b) DNA precipitation and purification
   (c) Selection of PCR primers
3. Selection of an adequate thermostable polymerase and PCR reaction buffer

The size of generated RAPD bands is related to the DNA extraction protocol, but there is a much stronger relationship between RAPD band sizes and relative PCR-DNA template concentration, i.e. volume of arDNA suspension applied in a PCR reaction and/or dilution of arDNA extract. This is probably related to the presence of PCR inhibitors in a particular PCR reaction which is again related to the DNA extraction protocol and pre-extraction treatments of tissue samples. The size of RAPD bands is also related to the RAPD primer and thermostable polymerase applied in a particular PCR amplification. The type of *Taq* polymerase and PCR reaction buffer significantly affected the sizes of generated RAPD-PCR fragments and their PCR yield (see Fig. 28; p 159).

## *3.2. Cloning and Sequence analysis*

RAPD results obtained from preserved specimens in this study cannot alone be used as reliable markers for *Nezumia*. Southern blotting analyses aided investigations on the homology of RAPD-PCR products, but more specific information on the genetic structure of *Nezumia aequalis* and *N. micronychodon* was still missing. RFLP investigations, cloning and sequencing of selected RAPD products were used to obtain more specific molecular information on the investigated species.

The success of cloning with selected RAPD-PCR DNA products was occasionally examined by PCR, but mainly by digesting recombinant pUC18 (with inserted RAPD-PCR product) with *EcoR* I and *Hind* III. A relatively large number of recombinant plasmids that contained the insert of correct size were produced (Fig. 34), but the transformed *E. coli* did not exhibit good growth in LB broth/ampicillin medium with all of them. This could be due to incompatibility between host cells and recombinant plasmid, medium, or other reasons. If routine checking indicated insert of the correct size and cells showed good growth, these recombinant plasmids were selected for DIG-labelling, Southern blotting, RFLP and/or sequencing experiments.

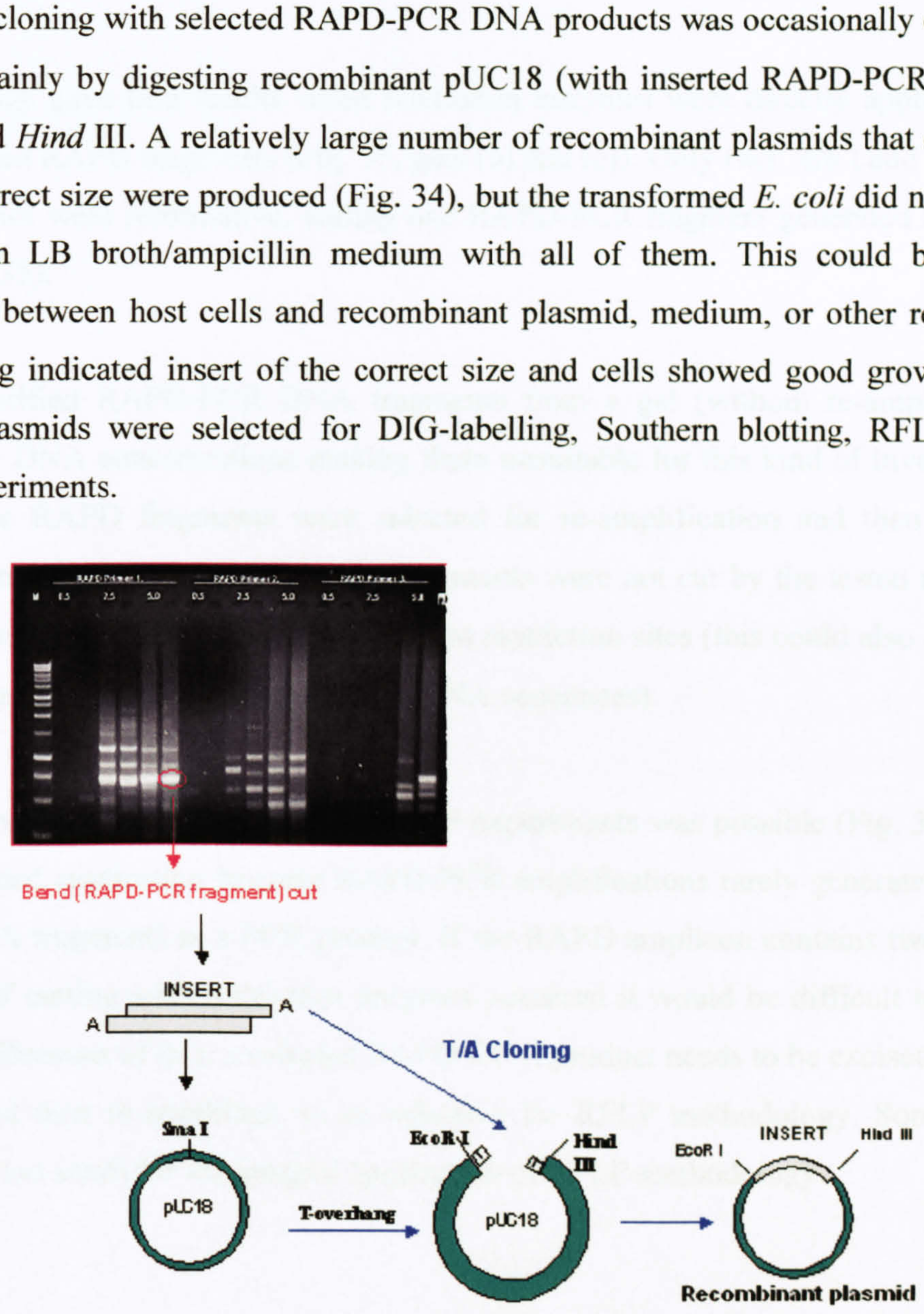



**Figure 34** – T/A cloning of RAPD-PCR DNA fragment.

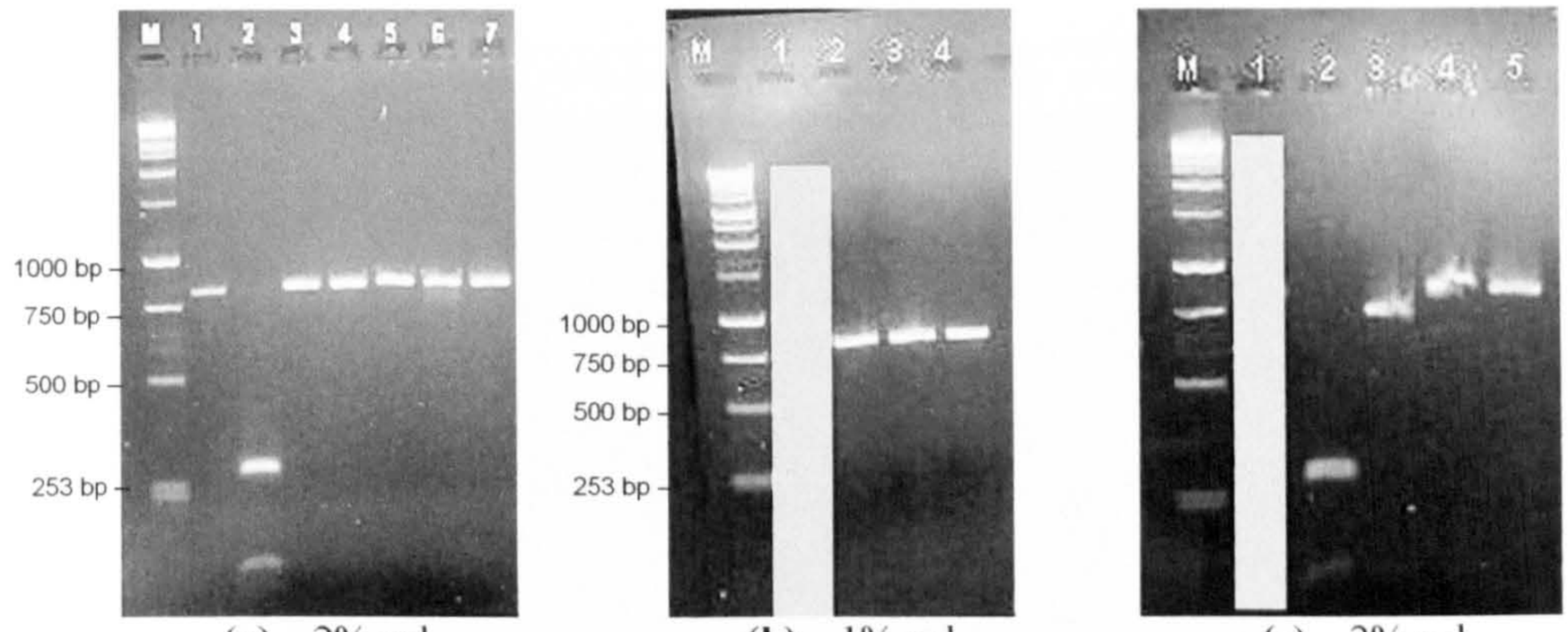
### 3.2.1. RFLP experiments

Total genomic DNA, RAPD-PCR products and cloned RAPD-PCR fragments were used for RFLP experiments. As expected, total genomic DNA extracted from formalin preserved tissue proved to be unsuitable for this kind of investigation because of arDNA degradation and inability to visualise fragments of arDNA on the agarose gel. As mentioned before, formalin-arDNA was so degraded that it could not be visualised even on 2% or 3% gels. For the same reason, it was not possible to visualise RFLP fragments if restriction nucleases were applied to total genomic DNA.

This methodology gave best results when restriction enzymes were directly applied to the "liberated" cloned RAPD fragments (Fig. 35, gels (b) and (c)). Only two, *Alu* I and *Ava* II, of all tested enzymes were informative, cutting one RAPD-PCR fragment generated by RAPD primer 44 (Fig. 35).

Excised and purified RAPD-PCR DNA fragments from a gel (without re-amplification) usually had low DNA concentrations making them unsuitable for this kind of investigation. Therefore, some RAPD fragments were selected for re-amplification and then used for restriction enzyme digestion. Most of these fragments were not cut by the tested restriction enzymes presumably because they did not contain restriction sites (this could also be caused by damage to enzyme restriction sites in the arDNA sequences).

Direct use of unpurified PCR products for RFLP experiments was possible (Fig. 35(a)), but this was of limited application because RAPD-PCR amplifications rarely generate only one band (PCR-DNA fragment) as a PCR product. If the RAPD amplicon contains two or more products, then if cutting with restriction enzymes occurred it would be difficult to identify which was cut. Because of this, a selected RAPD-DNA product needs to be excised from the gel, purified and then re-amplified, to be effective for RFLP methodology. Some RAPD fragments were too small for meaningful application of RFLP methodology.




**(a)** – 2% gel **(b)** – 1% gel **(c)** – 2% gel

**Figure 35 – RFLP experiments. Gel (a) -** with RAPD-PCR products; and **gels (b) and (c)** – with cloned, "liberated", RAPD-PCR fragments. ***Alu* I** cut cloned RAPD fragments ((c)-L: 2) and PCR product size of ~800 bp ((a)-L: 2). The ***Nde* I** ((c)-L: 3) may also cut one of the cloned fragment, but fragments were so faint; the obtained results were not clear and informative (the experiment was repeated, but results were similar).
"M" indicates size marker fragments – 1 Kb Ladder (Promega)

All this led to unreliable application of RFLP methodology to PCR products in the study. This kind of information on the investigated specimens was of little use in obtaining specific molecular data on the investigated species, or for development of other molecular markers. The initial results suggested that it was not worthwhile to continue this kind of investigation and so further investigations were redirected to sequencing recombinant plasmids and developing specific PCR primers for *N. aequalis* and *N. micronychodon* to obtain more specific information on their sequences.

### 3.2.2. Sequencing of cloned RAPD fragments and primers development

Four recombinant plasmids (pUC18 with inserted RAPD-PCR fragment) were sequenced (Fig. 36).

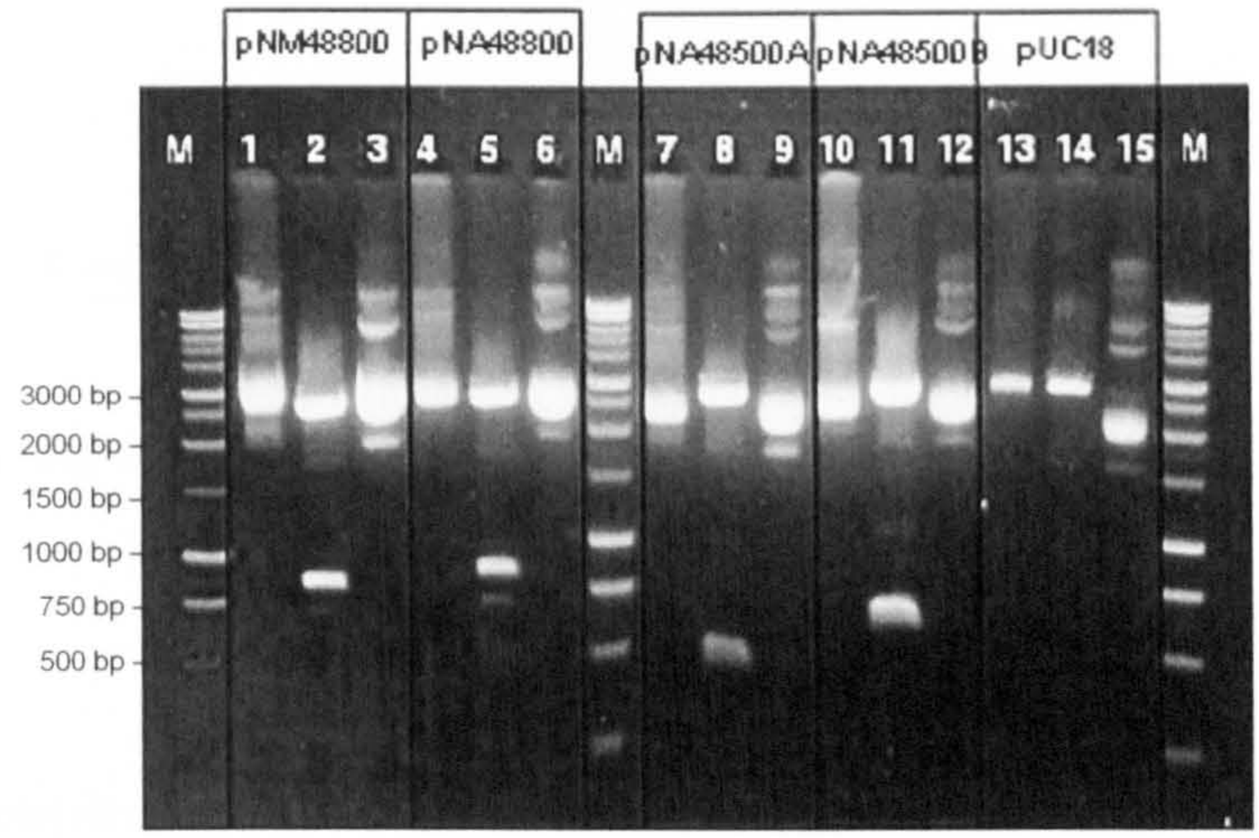



**Figure 36** – Recombinant plasmids used for sequencing of cloned RAPD fragments generated by RAPD Primer 44. "M" indicates size marker fragments – 1 Kb DNA Ladder (Promega).

**Lanes 1 - 3:** recombinant pNM48800 containing RAPD-PCR fragment size of ~800 bp generated with *Nezumia micronychodon* formalin-arDNA using RAPD primer 44.
**Lane 1:** pNM48800 cut with *Sma* I
**Lane 2:** pNM48800 cut with *EcoR* I and *Hind* III
**Lane 3:** uncut (non-digested) pNM48800

**Lanes 4 - 6:** recombinant pNA48800 containing RAPD-PCR fragment size of ~800 bp generated with *Nezumia aequalis* formalin-arDNA using RAPD primer 44.
**Lane 4:** pNA48800 cut with *Sma* I
**Lane 5:** pNA48800 cut with *EcoR* I and *Hind* III
**Lane 6:** uncut (non-digested) pNA48800

**Lanes 7 - 9:** recombinant pNA48500A containing RAPD-PCR fragment size of ~500 bp generated with *Nezumia aequalis* formalin-arDNA using RAPD primer 44.
**Lane 7:** pNA48500A cut with *Sma* I
**Lane 8:** pNA48500A cut with *EcoR* I and *Hind* III
**Lane 9:** uncut (non-digested) pNA48500A

**Lanes 10 - 12:** recombinant pNA48500B containing RAPD-PCR fragment size of ~500 bp generated with *Nezumia aequalis* formalin-arDNA using RAPD primer 44.
**Lane 10:** pNA48500B cut with *Sma* I
**Lane 11:** pNA48500B cut with *EcoR* I and *Hind* III
**Lane 12:** uncut (non-digested) pNA48500B

**Lanes 13 - 15:** unmodified pUC18;
**Lane 13:** pUC18 cut with *Sma* I
**Lane 14:** pUC18 cut with *EcoR* I and *Hind* III
**Lane 15:** uncut (non-digested) pUC18

In sequenced strands, it was expected to see the sequence of the cloned RAPD fragment (with forward and reverse-complement RAPD primer, "T" from T-overhang, and "A" added by *Taq* terminal transferase activity during RAPD-PCR) between "ccc" and "ggg" of *Sma* I, i.e. between *Hind* III and *EcoR* I restriction sites of the vector (Fig. 37).

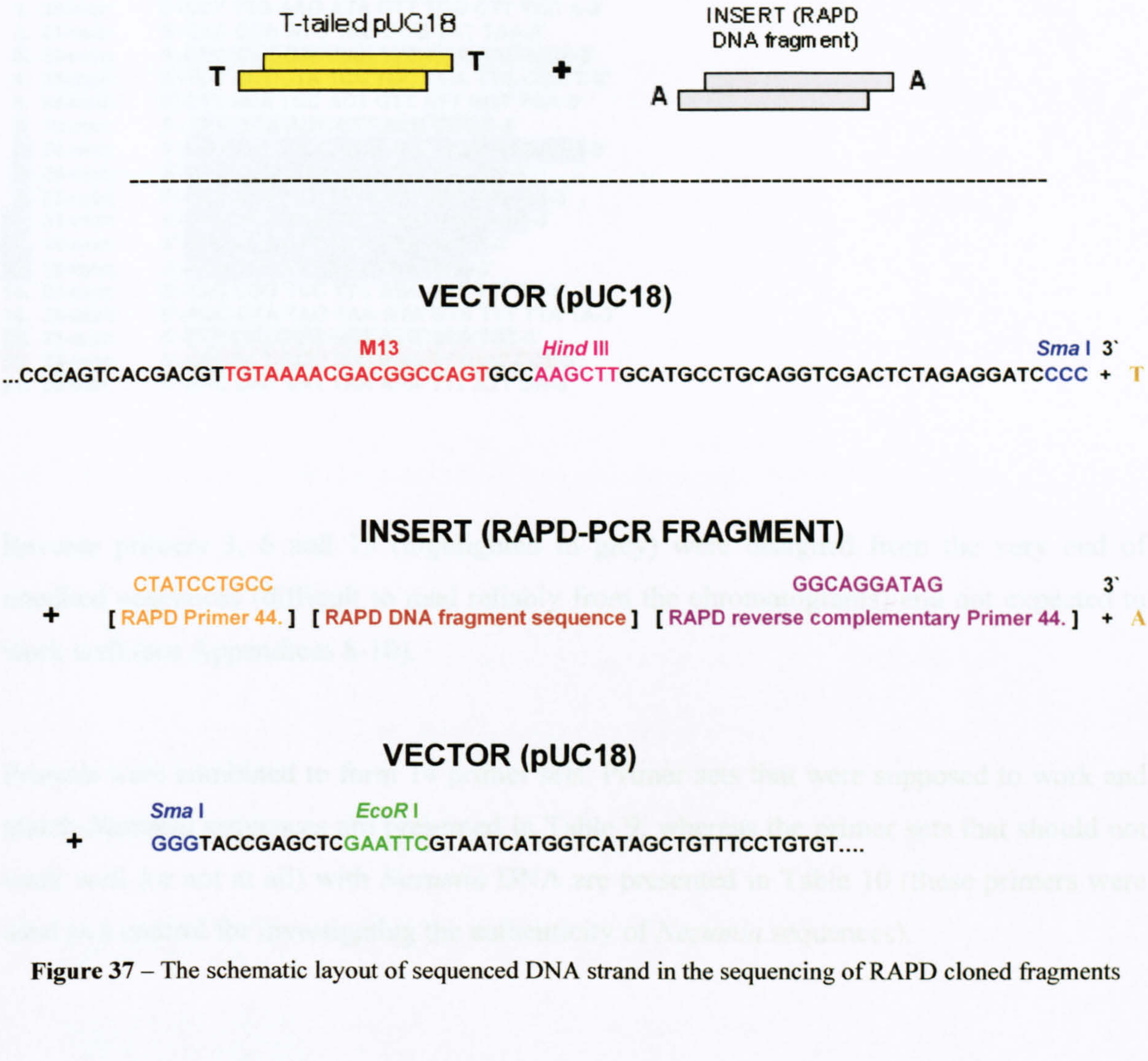



**Figure 37** – The schematic layout of sequenced DNA strand in the sequencing of RAPD cloned fragments

The raw sequencing data (unedited) with chromatograms and basic layout of sequenced strands are shown in Appendices 7-10. These unedited sequences were used to design specific primers for *N. aequalis* and *N. micronychodon*. Primer design was carried out "by eye" and applied criteria were to the length of primers (to be long at least 18 bases) and to the possibility of generating PCR products of different sizes. Two categories of specific primers were designed, one that should work well and match the sequence of *Nezumia*, and

another that does not work well (or not at all) with a genuine DNA of *Nezumia* because sequences of primers are partly (highlighted in yellow) or completely (highlighted in red) in the sequence of the vector, or they are designed from the artifactual sequence of RAPD primer repeats (the sequence of these primers is highlighted in blue; see Appendix 7). In total, seventeen primers, length 18 - 26 basis, were designed:

| # | | *Sequences of primers* |
|---|---|---|
| 1. | 25-mer: | 5`-CCT TTG AAG ATA GTT TGG CTT TAC A-3` |
| 2. | 21-mer: | 5`-CAC CCA GTG TAG CGG TAT TAA-3` |
| 3. | 24-mer: | 5`-CCA AAC CAT TAA TTC CGA AAG GCA-3` |
| 4. | 25-mer: | 5`-TCC CCT CTA TCC TGC CGA TTG GGA T-3` |
| 5. | 24-mer: | 5`-CTT ACA TCC ACT GTT ATT GGT TGA-3` |
| 6. | 19-mer: | 5`- GGT TTA ACA CTT AAG GCT T-3` |
| 7. | 24-mer: | 5`-GGA TAG GGC AGG ATA GGG CAG GAT-3` |
| 8. | 20-mer: | 5`-AGC ACG GTA CCC TCT ATC CT-3` |
| 9. | 23-mer: | 5`-AAA ATG TTT TTT CCT GGG GTA AT-3` |
| 10. | 21-mer: | 5`-TAG GCA GGA TCT GGC AGG ATA-3` |
| 11. | 19-mer: | 5`-ATC CTG CAT ATA TCC ATG C-3` |
| 12. | 18-mer: | 5`-AAA TTC ATT GAA TCA AGA-3` |
| 13. | 23-mer: | 5`-TAC CGG TCC TTC AGC TAT TTT TT-3` |
| 14. | 26-mer: | 5`-AGC GTA TAG TAA GTA GTA TTT TTA TA-3` |
| 15. | 21-mer: | 5`-TTT TAC CGG GGT ATG GAA TAT-3` |
| 18. | 23-mer: | 5`-CTT GCT GCC TGC AGG TCG ACT CT-3` |
| 21. | 23-mer: | 5`-AAC GAT CAT TAA AAA TTT ATT GA-3` |

Reverse primers 3, 6 and 15 (highlighted in grey) were designed from the very end of unedited sequences (difficult to read reliably from the chromatograms) and not expected to work well (see Appendices 8-10).

Primers were combined to form 14 primer sets. Primer sets that were supposed to work and match *Nezumia* sequences are presented in Table 9, whereas the primer sets that should not work well (or not at all) with *Nezumia* DNA are presented in Table 10 (these primers were used as a control for investigating the authenticity of *Nezumia* sequences).

**Table 9** - List of *Nezumia*-specific (RAPD-PCR derived) primer sets that should work with *Nezumia* DNA

| *Set* | Name and size of primers | *Sequence of primers* | *PCR product size* |
|---|---|---|---|
| *8)* | 1. NA108800F2, 25-mer: | 5`-CCT TTG AAG ATA GTT TGG CTT TAC A-3` | |
| | 2. NA108800R1, 21-mer: | 5`-CAC CCA GTG TAG CGG TAT TAA-3` | 350 bp |
| *9)* | 1. NA108800F2, 25-mer: | 5`-CCT TTG AAG ATA GTT TGG CTT TAC A-3` | |
| | 3. NA108800R2, 24-mer: | 5`-CCA AAC CAT TAA TTC CGA AAG GCA-3` | 500 bp |
| *16)* | 21. NM48800F1, 23-mer: | 5`-AAC GAT CAT TAA AAA TTT ATT GA-3` | |
| | 14. NM48800 R1, 26-mer: | 5`-AGC GTA TAG TAA GTA GTA TTT TTA TA-3` | 350 bp |
| *17)* | 21. NM48800F1, 23-mer: | 5`-AAC GAT CAT TAA AAA TTT ATT GA-3` | |
| | 15. NM48800 R2, 21-mer: | 5`-TTT TAC CGG GGT ATG GAA TAT-3` | 500 bp |
| *18)* | 13. NM48800 F2, 23-mer: | 5`-TAC CGG TCC TTC AGC TAT TTT TT-3` | |
| | 14. NM48800 R1, 26-mer: | 5`-AGC GTA TAG TAA GTA GTA TTT TTA TA-3` | 300 bp |
| *19)* | 13. NM48800 F2, 23-mer: | 5`-TAC CGG TCC TTC AGC TAT TTT TT-3` | |
| | 15. NM48800 R2, 21-mer: | 5`-TTT TAC CGG GGT ATG GAA TAT-3` | 450 bp |

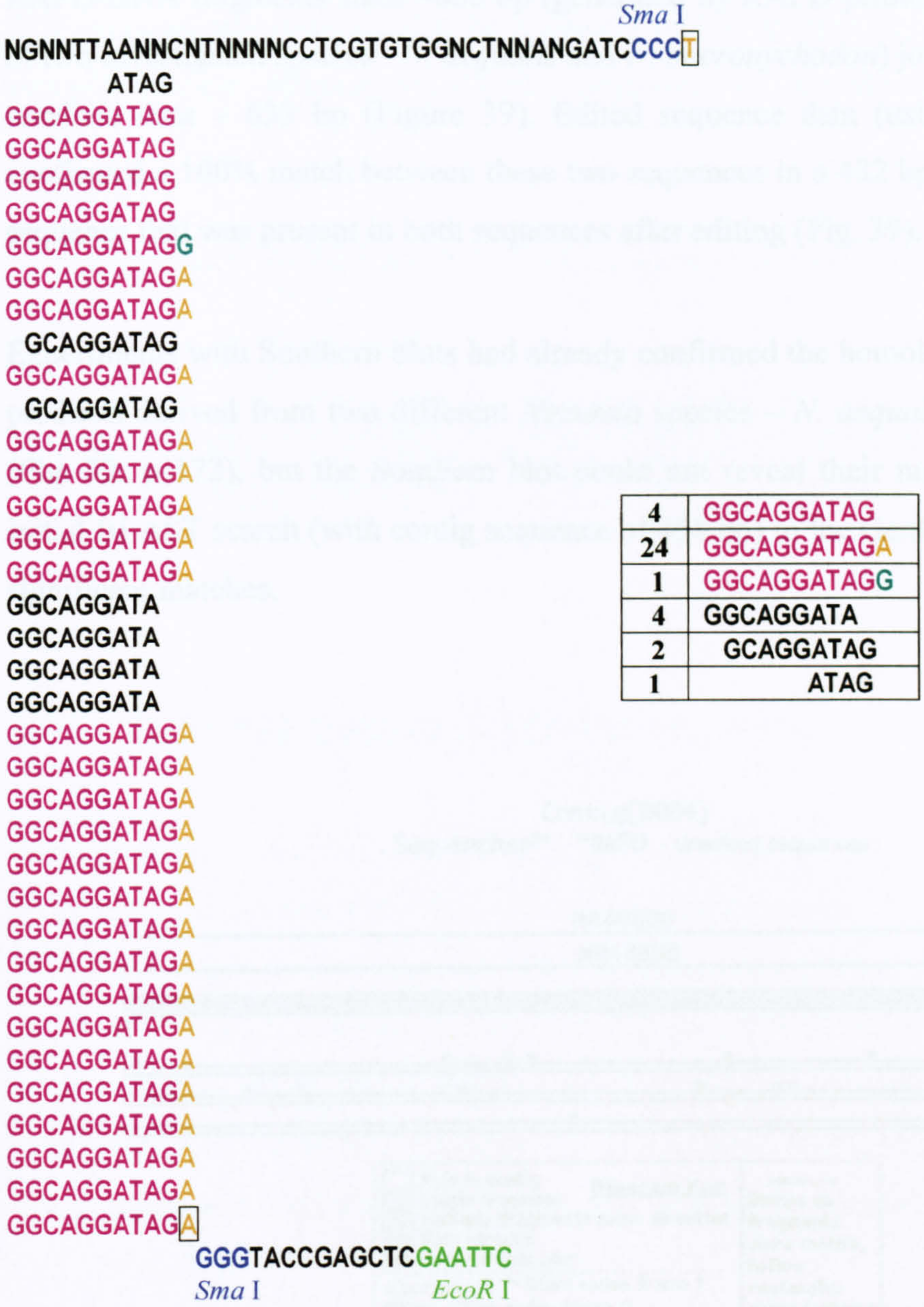
**Table 10** - List of specific primer sets that should not work well (or not at all) with *Nezumia* DNA

| *Set* | Name and size of primers | *Sequence of primers* | *PCR product size* |
|---|---|---|---|
| *6)* | 18. NA108800 F1, 23-mer: | 5`- CTT GCT GCC TGC AGG TCG ACT CT-3` | |
| | 2. NA108800R1, 21-mer: | 5`-CAC CCA GTG TAG CGG TAT TAA-3` | 400-450 bp |
| *7)* | 18. NA108800 F1, 23-mer: | 5`- CTT GCT GCC TGC AGG TCG ACT CT-3` | |
| | 3. NA108800R2, 24-mer: | 5`-CCA AAC CAT TAA TTC CGA AAG GCA-3` | 600 bp |
| *10)* | 4. NA48500BF, 25-mer: | 5`-TCC CCT CTA TCC TGC CGA TTG GGA T-3` | |
| | 5. NA48500BR1, 24-mer: | 5`-CTT ACA TCC ACT GTT ATT GGT TGA-3` | 350 bp |
| *11)* | 4. NA48500BF, 25-mer: | 5`-TCC CCT CTA TCC TGC CGA TTG GGA T-3` | |
| | 6. NA48500BR2, 19-mer: | 5`- GGT TTA ACA CTT AAG GCT T-3` | 550 bp |
| *12)* | 7. NA48500A2F, 24-mer: | 5`-GGA TAG GGC AGG ATA GGG CAG GAT-3` | |
| | 8. NA48500A2R1, 20-mer: | 5`-AGC ACG GTA CCC TCT ATC CT-3` | 350 bp |
| *13)* | 7. NA48500A2F, 24-mer: | 5`-GGA TAG GGC AGG ATA GGG CAG GAT-3` | |
| | 9. NA48500A2R2, 23-mer: | 5`-AAA ATG TTT TTT CCT GGG GTA AT-3` | 700 bp |
| *14)* | 10. NA48500A1F, 21-mer: | 5`-TAG GCA GGA TCT GGC AGG ATA-3` | |
| | 11. NA48500A1R1, 19-mer: | 5`-ATC CTG CAT ATA TCC ATG C-3` | 305 bp |
| *15)* | 10. NA48500A1F, 21-mer: | 5`-TAG GCA GGA TCT GGC AGG ATA-3` | |
| | 12. NA48500A1R2, 18-mer: | 5`-AAA TTC ATT GAA TCA AGA-3` | 570 bp |

The first attempt at sequencing NA 48500A insert gave a sequence of poor quality (unclear sequence with strong background), which was difficult to read and edit with confidence (Appendix 7 (a)). The sequencing was repeated with the same sample (recombinant plasmid pNA48500A). This provided a sequence of good quality, clear and easy to read from the chromatogram (Appendix 7 (b)). Repeats of reverse and complementary RAPD primer 44 (RAPD primer used for generating this particular RAPD-PCR fragment) were observed (Fig. 38). The reverse and complementary RAPD primer 44 in this 372 bp-long "insert" sequence was repeated 36 times (including complete and incomplete RAPD primer):

| | | |
|---|---|---|
| **4** | GGCAGGATAG | **- complete rev. and comp. RAPD primer 44** |
| **24** | GGCAGGATAGA | **- additional "A" at the 3` end** |
| **1** | GGCAGGATAGG | **- additional "G" at the 3` end** |
| **4** | GGCAGGATA | **- incomplete rev. and comp. RAPD primer 44** |
| **2** | GCAGGATAG | **- incomplete rev. and comp. RAPD primer 44** |
| **1** | ATAG | **- incomplete rev. and comp. RAPD primer 44** |

The sequence of NA48500A insert appears to be a PCR/cloning artifact. Primers derived from this NA48500A sequence with repeats (primer sets: 12, 13, 14 and 15) did not produce successful amplifications with arDNA of *Nezumia aequalis* and *N. micronychodon* (Appendix 12), confirming that this sequence is artifactual.


Sma I
NGNNTTAANNCNTNNNNCCTCGTGTGGNCTNNANGATCCCCT
ATAG
GGCAGGATAG
GGCAGGATAG
GGCAGGATAG
GGCAGGATAG
GGCAGGATAGG
GGCAGGATAGA
GGCAGGATAGA
GCAGGATAG
GGCAGGATAGA
GGCAGGATAGA
GGCAGGATAGA
GGCAGGATAGA
GGCAGGATAGA
GGCAGGATA
GGCAGGATA
GGCAGGATA
GGCAGGATA
GGCAGGATAGA
GGCAGGATAGA
GGCAGGATAGA
GGCAGGATAG

The NM48800 (Appendix 9) and NA48800 (Appendix 10) insert sequences of cloned RAPD-DNA fragments sizes ~800 bp (generated by RAPD primer 44 and formalin-arDNA of two investigated species - *N. aequalis* and *N. micronychodon*) joined in one contig even as unedited data – 633 bp (Figure 39). Edited sequence data (using Sequencher software) confirmed a 100% match between these two sequences in a 432 bp segment - the part of the sequence that was present in both sequences after editing (Fig. 39).

Experiments with Southern blots had already confirmed the homology of these RAPD-DNA products derived from two different *Nezumia* species – *N. aequalis* and *N. micronychodon* (Fig. 32; p 172), but the Southern blot could not reveal their nucleic acid sequences. An initial BLAST search (with contig sequence of 458 bp) in the GenBank database detected no significant matches.

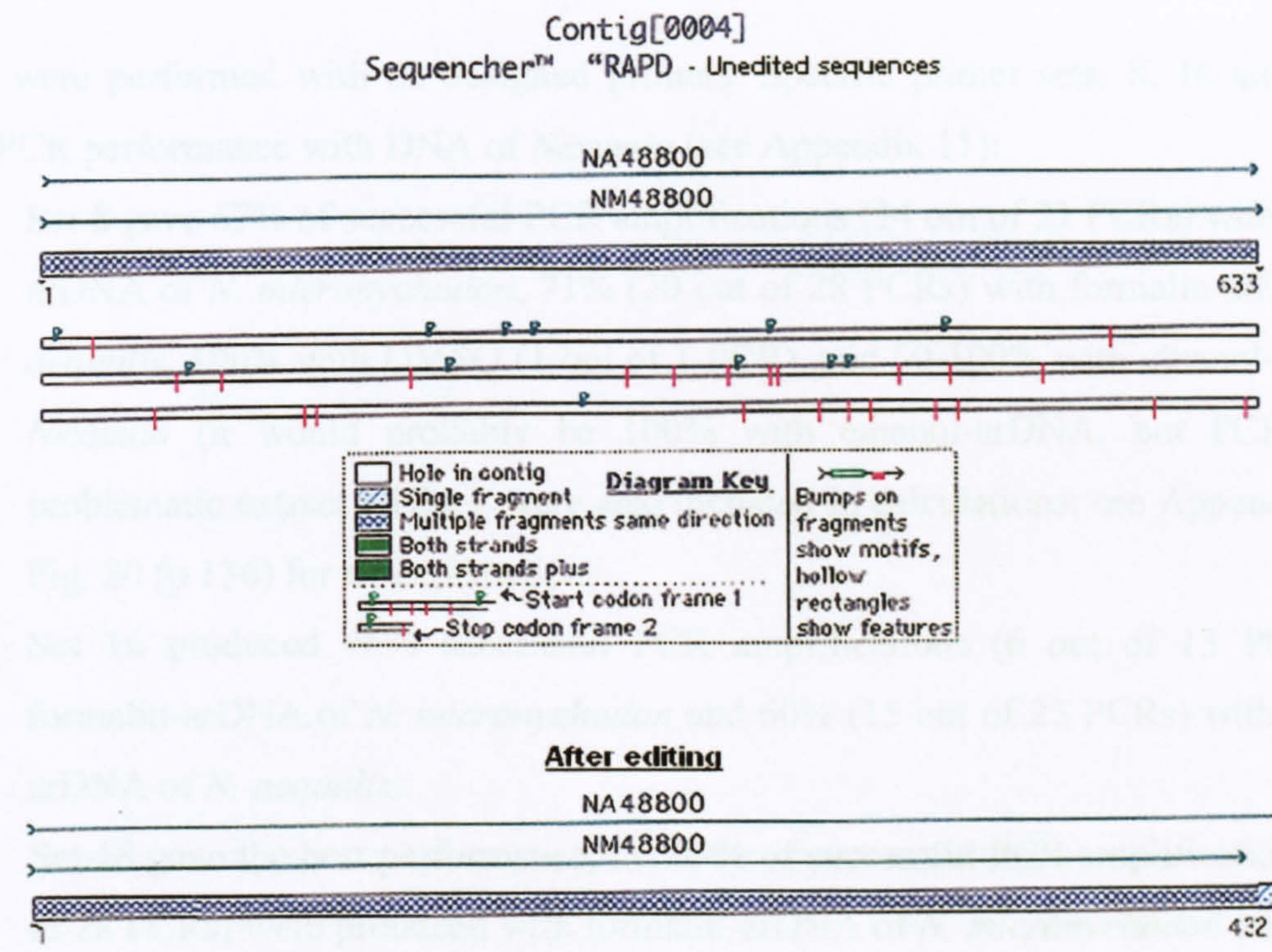



**Figure 39** – The match of NA48800 and NM48800 sequences (before and after editing, i.e. before and after removing the residues of the pUC18 vector sequence). These sequences were obtained by cloning RAPD-PCR DNA fragments size of ~800 bp generated by RAPD primer 44 and formalin-arDNA of *Nezumia aequalis* (NA48800) and *N. micronychodon* (NM48800).

Sequences of four cloned RAPD fragments were aligned against each other, but only these two sequences (NA 48800 and NM 48800) joined in one contig. The NA 48500B sequence (358 bp-long) was considered to be a sequence from another (anonymous) region of the *N. aequalis* genome, whereas NA 48500A sequence was considered as being a PCR/cloning artifact.

Both sequences (NA48800 and NA 48500B) are A, T rich regions:

| | *A* | *T* | *C* | *G* | *A,T* | *C,G* |
|---|---|---|---|---|---|---|
| **NA 48800** (458 bp) | 121 (26%) | 166 (36%) | 85 (19%) | 86 (19%) | 287 (62%) | 171 (38%) |
| **NA 48500B** (358 bp) | 108 (30%) | 117 (33%) | 47 (13%) | 84 (23%) | 225 (63%) | 131 (37%) |

### 3.2.3. PCRs with designed (RAPD-derived) primers

PCRs were performed with all designed primers. Specific primer sets: 8, 16 and 18 gave good PCR performance with DNA of *Nezumia* (see Appendix 11):

- **Set 8** gave 67% of successful PCR amplifications (14 out of 21 PCRs) with formalin-arDNA of *N. micronychodon*, 71% (20 out of 28 PCRs) with formalin-arDNA of *N. aequalis*, 100% with DMSO (1 out of 1 PCR), and 50-100% with ethanol-arDNA of *Nezumia* (it would probably be 100% with ethanol-arDNA, but PCRs with a problematic extract NHM-2 were also included in calculations; see Appendix 11 and Fig. 20 (p 136) for an explanation).
- **Set 16** produced 46% successful PCR amplifications (6 out of 13 PCRs) with formalin-arDNA of *N. micronychodon* and 60% (15 out of 25 PCRs) with formalin-arDNA of *N. aequalis*.
- **Set 18** gave the best performance, i.e. 89% of successful PCR amplifications (25 out of 28 PCRs) were produced with formalin-arDNA of *N. micronychodon* and 94% (34 out of 36 PCRs) with formalin-arDNA of *N. aequalis*.

These three primers are designed to amplify the same (unknown) region of *Nezumia* DNA genome (amplicons sizes of 300-350 bp), but their amplification efficiency was different. This emphasises the importance of designing primers, i.e. testing different "ready primers",

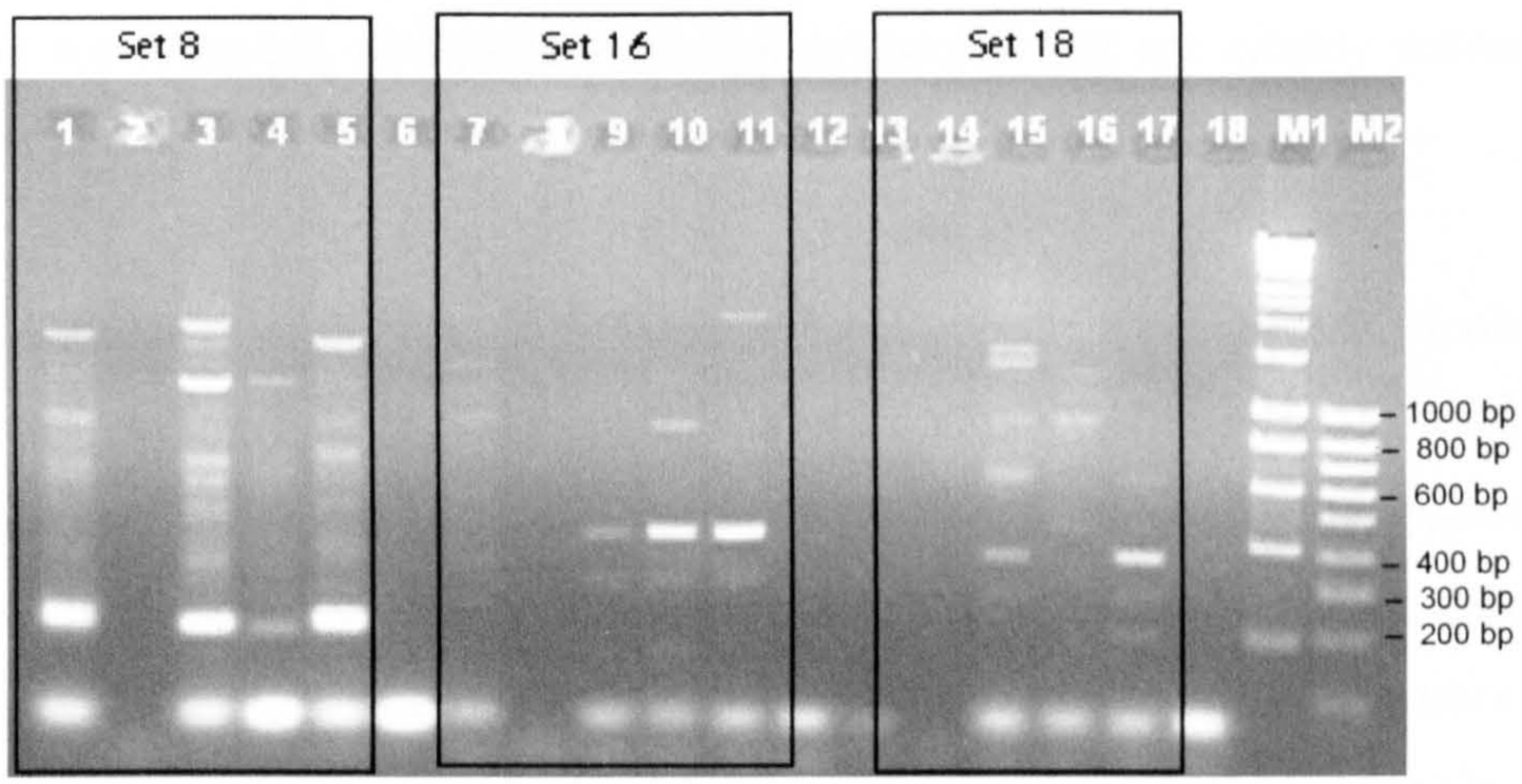
for the best amplification efficiency with preserved specimens. In general, these three primer sets were efficient in amplifying formalin-Steedman-arDNA of *Nezumia*, but primer set 18 had the best amplification efficiency for this region (see Appendix 11). Primer sets 8, 16 and 18 also gave some successful PCR amplifications with DNA from control, ethanol preserved specimens of *Coryphaenoides armatus* and *C. rupestris* (species closely related to *Nezumia*), but not with DNA of cod, rainbow trout and mackerel – tested on a limited number of samples. The results indicate the possibility of generating PCR products in two *Nezumia* and closely related species (*Coryphaenoides* species), but not in three other, more distant species.

Other primer sets (9, 17 and 19) gave poor PCR performances (see Appendix 11). For example, primer set 9 gave 9% of successful amplifications (3 out of 34 PCRs) with formalin-arDNA of *Nezumia,* whereas primer set 17 did not generate PCR products with any of tested formalin arDNA. This is not surprising because some of primers (primers 3 and 15) that formed these three primer sets were designed from regions of sequences that were difficult to read from the chromatograms, and therefore, the base compositions of these primers might not exactly match priming sites in DNA sequences of *Nezumia* (see Appendix 9 and Table 9). However, surprisingly, primer sets 6, 10 and 11 gave a few successful amplifications with formalin-arDNA of *N. micronychodon* and *N. aequalis*, and control ethanol-DNA of *Coryphaenoides* as well (Appendix 12). What exactly caused these successful amplifications with *Nezumia* and control specimens of *Coryphaenoides* is unclear.

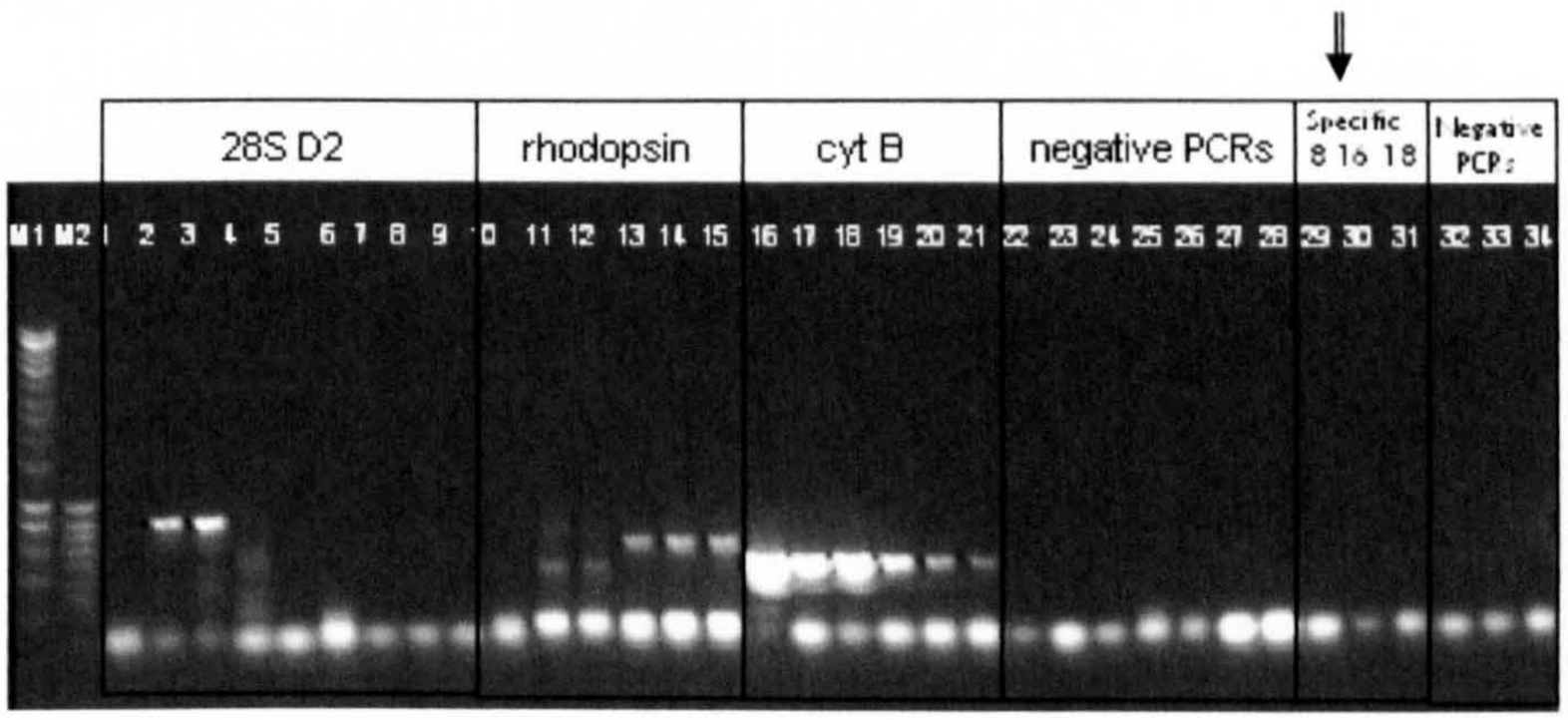
### 3.2.4. PCR validation, sequencing and sequence validation

Besides PCR validation on arDNA from the DMSO-preserved tissue sample of *Nezumia* cf. *aequalis* that came from Greenland (experiments performed in the Nescot`s Molecular Laboratory, the laboratory that was normally used for the project experiments) - Fig. 41(c)-Lanes: 8 and 11), PCRs with *Nezumia*-specific sets 8, 16 and 18 were also validated in the Molecular Laboratory of the Natural History Museum (NHM), London (Fig. 40, gel (a)). These PCR experiments were performed on ethanol preserved samples of *N. micronychodon* and *N. aequalis*, tissue samples from fish specimens that came from different locations and collections (see General material and methods). Ethanol-arDNAs were extracted by applying protocol (H), i.e. using a different DNA extraction protocol than was applied for production of DMSO- and formalin-Steedman's-arDNA extracts in the Nescot laboratory. Also, a

different *Taq* polymerase, a different type of PCR thermocycler, and other PCR conditions were applied in the NHM laboratory. A PCR product generated by primer set 18 was subjected to direct sequencing and used for sequence validation (Fig. 43).




**(a) - *Nezumia***




**(b) – mackerel *(Scomber scombrus)***

**Figure 40 – (a)** PCRs with designed *Nezumia*-specific primer sets: 8, 16 and 18 using control ethanol-arDNA of *Nezumia aequalis* and *N. micronychodon*. All tested primers produced PCR products with tested DNA extractions except with the extract no. NHM-2 (Lanes: 2, 8, 14) which proved to be problematic - probably contained strong PCR inhibitors (see section 3.1.1.2. and Fig. 20, p 136 for an explanation). Further PCR optimisation was required in order to generate only one band of appropriate sizes.

**(b)** *Nezumia*-specific (RAPD-derived) primer sets: 8, 16 and 18 with DNA of mackerel did not generate any PCR products (Lanes: 29-31, indicated by arrow). Other applied specific primers produced PCR products with the same DNA extract of mackerel (Lanes: 1-9 for 28S rDNA, Lanes: 10-15 for rhodopsin, and Lanes: 16-21 for cytochrome b).
"M1" indicates quantitative size marker fragments – HyperLadder I (Bioline)
"M2" indicates quantitative size marker fragments – HyperLadder IV (Bioline)

Primer sets 8, 16 and 18 were tested against DNA of fresh/frozen mackerel (*Scomber scombrus*; family Scombridae). None of these primers generated PCR product (Fig. 40 (b)-L: 29-31). The same DNA extract of mackerel was used successfully for other PCRs with different marker systems (Fig. 40 (b)-L: 1-9 with 28S, L: 10-15 with rhodopsin, L: 16-21 with cyt b). *Scomber* species are taxonomically and evolutionarily distant from *Nezumia* (and the Family Macrouridae) which might suggest that designed primers 8, 16 and 18 are indeed specific for *Nezumia,* and perhaps for other species that are closely related to the genus *Nezumia* (see previous section and Appendix 11).

PCR products generated by the use of different *Nezumia*-specific primer sets (derived from the sequences of cloned RAPD fragments), with formalin-arDNA extracts isolated from different tissue types of one and different preserved specimens, were directly sequenced and aligned against the sequences of the cloned RAPD-PCR fragments in order to confirm the reproducibility of PCR amplifications and sequences (Figs. 41-44). The sequences of PCR products from DMSO and ethanol preserved samples of *Nezumia aequalis* were also aligned with sequences of the PCR products from formalin-fixed samples and cloned RAPD-DNA fragments (Fig. 44).

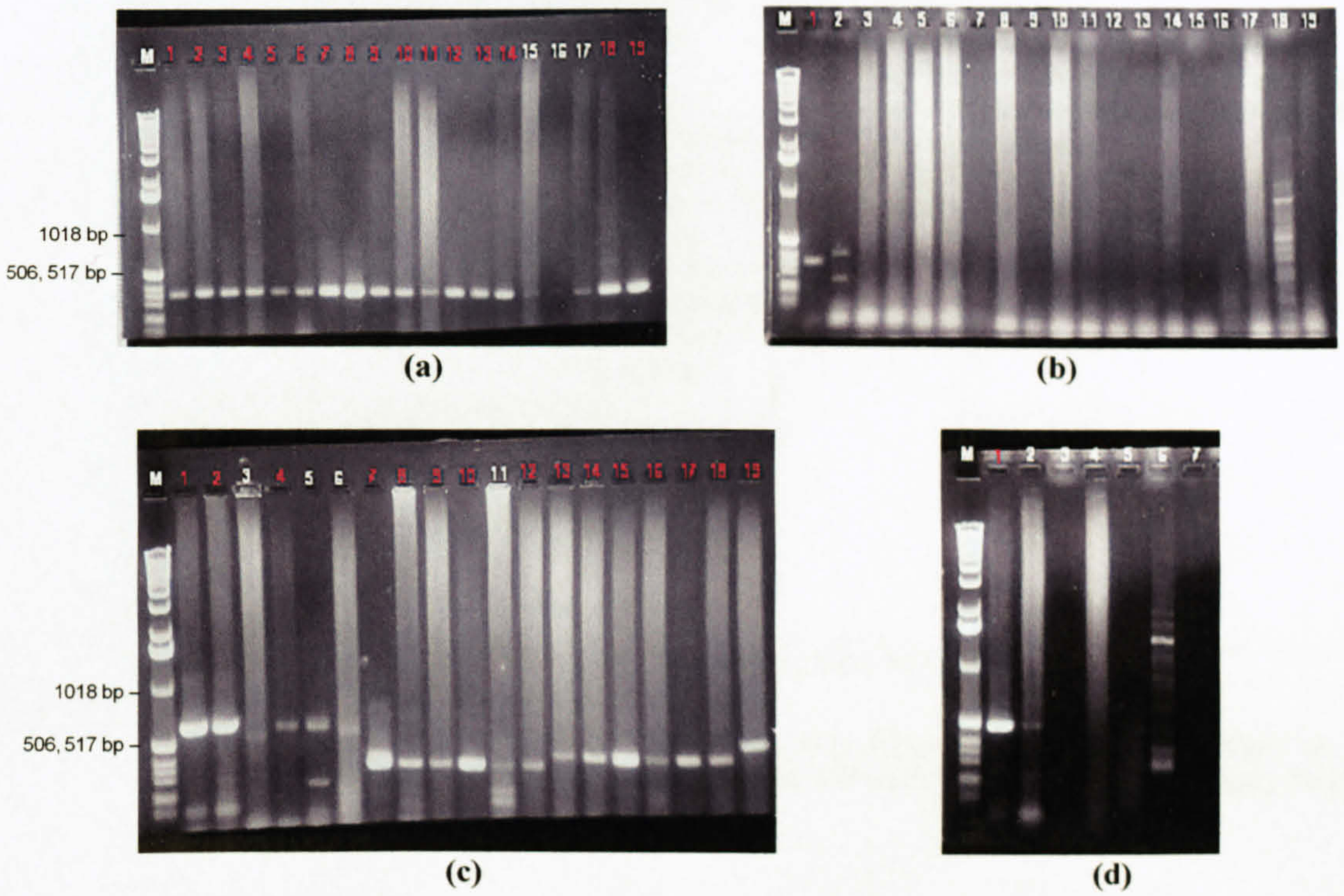



**Figure 41** – The first set of 32 PCRs sent to the MWG-Biotech for direct sequencing. The PCR products were generated with designed, RAPD-derived, specific primers and 16S mitochondrial PCR primer set. The lanes labelled in red indicate the PCR products sent for sequencing.
"M" indicates size marker fragments - 1 Kb DNA Ladder (Gibco).

**Gel (a):**

**Lane 1:** S1M18, *Set 18*, N.m R-5, Ex. 222 (formalin, muscle)
**Lane 2:** S2M18, *Set 18*, N.m R-13, Ex. 223 (formalin, muscle)
**Lane 3:** S3M18, *Set 18*, N.m R-10, Ex. 227 (formalin, muscle)
**Lane 4:** S4M18, *Set 18*, N.m R-10, Ex. 236 (formalin, eye)
**Lane 5:** S5M18, *Set 18*, N.m R-10, Ex. 237 (formalin, liver)
**Lane 6:** S6M18, *Set 18*, N.m R-5, Ex. 302 (formalin, vertebra with spine)
**Lane 7:** S8M18, *Set 18*, N.a R-23, Ex. 210 (formalin, muscle)
**Lane 8:** S9M18, *Set 18*, N.a R-8, Ex. 216 (formalin, muscle)
**Lane 9:** S10M18, *Set 18*, N.a R-7, Ex. 224 (formalin, muscle)
**Lane 10:** S11M18, *Set 18*, N.a R-25, Ex. 229 (formalin, muscle)
**Lane 11:** S12M18, *Set 18*, N.a R-26, Ex. 238 (formalin, muscle)
**Lane 12:** S13M18, *Set 18*, N.a R-20, Ex. 243 (formalin, muscle)
**Lane 13:** S14M18, *Set 18*, N.a R-24, Ex. 244 (formalin, muscle)
**Lane 14:** S15M18, *Set 18*, N.m R-27, Ex. 246 (formalin, mixed tissue – half part of body with head and all organs, but removed skin and intestines, and liver checked for the parasites presence)
**Lane 15:** *Set 8*, N.m R-10, Extr. 236 (formalin, eye)
**Lane 16:** *Set 8*, N.m R-10, Extr. 237 (formalin, liver)
**Lane 17:** *Set 8*, N.m R-5, Ex. 302 (formalin, vertebra with spine)
**Lane 18:** S21A8, *Set 8*, N.a R-23, Extr. 210 (formalin, muscle)
**Lane 19:** S22A8, *Set 8*, N.a R-8, Extr. 216 (formalin, muscle)

**Gel (b):**

**Lane 1:** S23A8, *Set 8*, N.a R-7, Ex. 224 (formalin, muscle)
**Lane 2:** *Set 10*, N.a R-8, Ex. 216 (formalin, muscle)
**Lane 3:** *Set 11*, N.m R-10, Ex. 227 (formalin, muscle)
**Lane 4:** *Set 11*, N.m R-12, Ex. 230 (formalin, muscle)
**Lane 5:** *Set 11*, N.m R-6, Ex. 231 (formalin, muscle)
**Lane 6:** *Set 11*, N.m R-15, Ex. 232 (formalin, muscle)
**Lane 7:** *Set 11*, N.m R-11, Ex. 233 (formalin, mixed tissue – almost a whole fish with all organs, but removed skin and intestines, and liver checked for the parasites presence)
**Lane 8:** *Set 11*, N.m R-10, Ex. 236 (formalin, eye)
**Lane 9:** *Set 11*, N.m R-10, Ex. 237 (formalin, liver)
**Lane 10:** *Set 11*, N.m R-16, Ex. 248 (formalin, muscle)
**Lane 11:** *Set 11*, N.m R-26, Ex. 302 (formalin, muscle
**Lane 12:** *Set 11*, N.a R-23, Ex. 210 (formalin, muscle)
**Lane 13:** *Set 11*, N.a R-8, Ex. 216 (formalin, muscle)
**Lane 14:** *Set 11*, N.a R-7, Ex. 224 (formalin, muscle)
**Lane 15:** *Set 11*, N.a R-28, Ex. 225 (formalin, muscle)
**Lane 16:** *Set 11*, N.a R-18, Ex. 228 (formalin, mixed tissue half part of body with head and all organs, but removed skin and intestines, and liver checked for the parasites presence)
**Lane 17:** *Set 11*, N.a R-25, Ex. 229 (formalin, muscle)
**Lane 18:** *Set 11*, C.r , Ex. $260^F$ (ethanol, muscle)
**Lane 19:** *Set 18*, "– " PCR, $H_2O$

**Gel (c):**

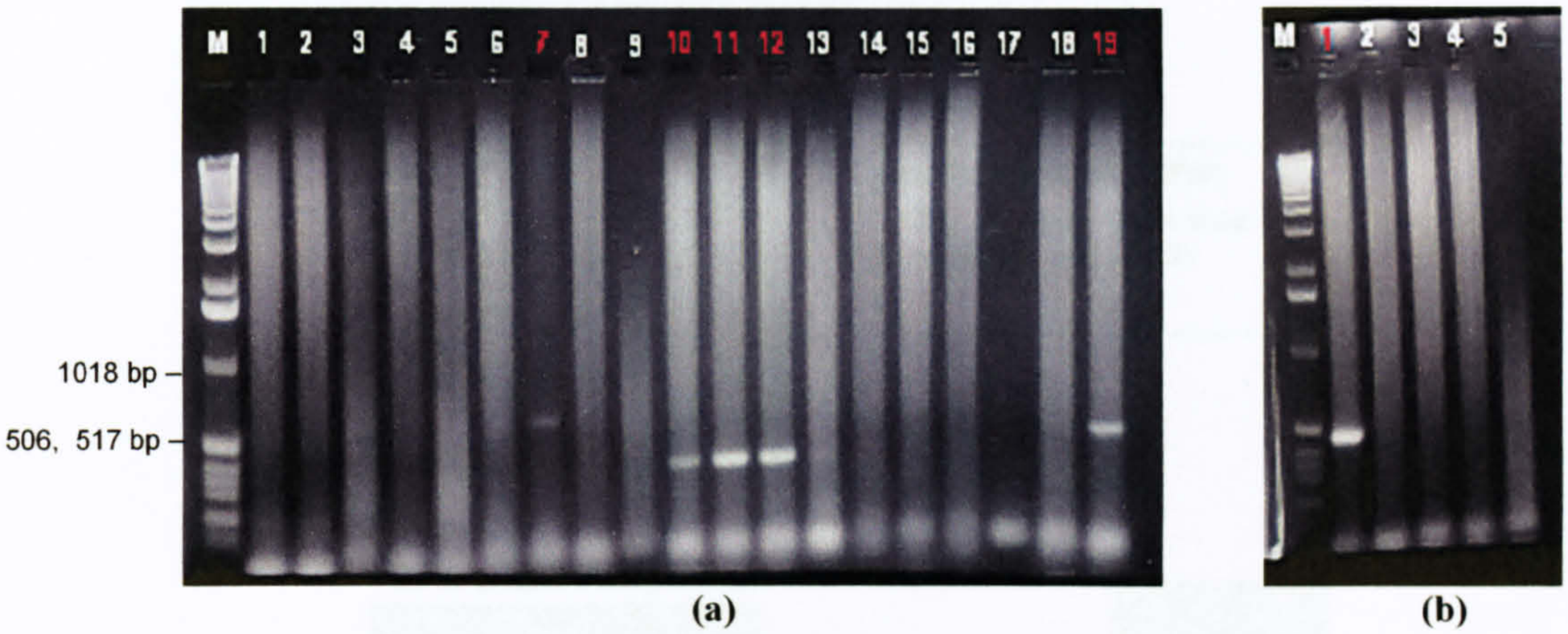
**Lane 1:** S18GM18, *mt-16S*, N.a (Greenl), Ex 300 (DMSO,muscl)
**Lane 2:** S19GM18, *mt-16S*, N.a (Greenl), Ex 300 (DMSO,muscl)
**Lane 3:** *mt-16S*, N.m R-14, Ex. 215 (formalin, muscle)
**Lane 4:** S20A18, *mt-16S*, N.a R-23, Ex. 210 (formalin, muscle)
**Lane 5:** *mt-16S*, N.a R-20, Ex. 243 (formalin, muscle)
**Lane 6:** *mt-16S*, N.a R-27, Ex. 246 (formalin, mixed tissue – half part of body with head and all organs, but removed skin and intestines, and liver checked for the the parasites presence)
**Lane 7:** S16A18, *Set 18*, N.a R-8, Ex. 216 (formalin, muscle)
**Lane 8:** S17G18, *Set 18*, ?N.a(Greenl),Ex. 301 (DMSO, muscle)
**Lane 9:** S7M18, *Set 18*, N.m R-14, Ex. 215 (formalin, muscle)
**Lane 10:** S24A8, *Set 8*, N.a R-8, Ex. 216 (formalin, muscle)
**Lane 11:** *Set 8*, N.a (Greenland), Ex. 301 (DMSO, muscle)
**Lane 12:** S25M8, *Set 8*, N.m R-14, Ex. 215 (formalin, muscle)
**Lane 13:** S26M16, *Set 16*, N.m R-14, Ex. 215 (formalin, muscle)
**Lane 14:** S27A16, *Set 16*, N.a R-23, Ex. 210 (formalin, muscle)
**Lane 15:** S28A16, *Set 16*, N.a R-8, Ex. 216 (formalin, muscle)
**Lane 16:** *Set 16*, N.a R-7, Ex. 224 (formalin, muscle)
**Lane 17:** S29A16, *Set 16*, N.a R-20, Ex. 243 (formalin, muscle)
**Lane 18:** S30A16, *Set 16*, N.a R-27, Ex. 246 (formalin, mixed tissue – half part of body with head and all organs, but removed skin and intestines, and liver checked for the parasites presence)
**Lane 19:** S31A6, *Set 6*, N.a R-23, Ex. 210 (formalin, muscle)

**Gel (d):**

**Lane 1:** S32A6, *Set 6*, N.a R-8, Ex. 216 (formalin, muscle)
**Lane 2:** *Set 6*, N.a R-20, Ex. 243 (formalin, muscle)
**Lane 3:** *Set 17*, N.a R-24, Ex. 244 (formalin, muscle)
**Lane 4:** *Set 17*, N.a R-27, Ex 246 (formalin, mixed tissue half part of body with head and all organs, but removed skin and intestines, and liver checked for the parasites presence)
**Lane 5:** *Set 17*, N.a R-31, Ex. 247a (formalin, mixed tissue half part of body with head and all organs, but removed skin and intestines, and liver checked for the parasites presence)
**Lane 6:** *Set 17*, C.r, Ex. $260^F$ (ethanol, muscle)
**Lane 7:** *Set 17*, "– " PCR, $H_2O$

**Ex.** – identifier number for DNA extract (e.g. Ex. 215 is extract no. 215)
**R** – identifier for fish individual (e.g., R-8 is fish no. 8)
**underlined** – PCR amplification sample sent for sequencing (e.g., S28A16 is amplification sample no. 28 (S28), which was generated with specific primer set 16 and with DNA of *Nezumia aequalis* (N.a))

Five additional PCR amplification samples (S40 - S45; Fig. 42) were sent for sequencing as a replacement for samples S18 – S25 (Fig. 41) that were accidentally destroyed by MWG-Biotech.




(a) (b)

**Figure 42** – Five PCR amplification reactions (replacement for the lost ones) sent additionally to MWG-Biotech for direct sequencing. PCR products were generated by using designed (RAPD-derived) primer sets 6 and 8, and 16S mitochondrial PCR primer set. The lanes labelled in red indicate the PCR products sent for sequencing.
"M" indicates size marker fragments - 1 Kb DNA Ladder (Gibco).

**Gel (a):**

**Lane 1:** *mt-16S*, N.a R-23, Ex. 210 (formalin, muscle)
**Lane 2:** *mt-16S*, N.a R-8, Ex. 216 (formalin, muscle)
**Lane 3:** *mt-16S*, N.a R-7, Ex. 224 (formalin, muscle)
**Lane 4:** *mt-16S*, N.a R-25, Ex. 229 (formalin, muscle)
**Lane 5:** *mt-16S*, N.m R-5, Ex. 222 (formalin, muscle)
**Lane 6:** *mt-16S*, N.m R-13, Ex. 223 (formalin, muscle)
**Lane 7:** S40MM, *mt-16S*, N.m R-5, Ex. 302 (formalin, vertebra with spine)
**Lane 8:** *mt-16S*, N.m R-14, Ex. 215 (formalin, muscle)
**Lane 9:** *mt-16S*, "– " PCR, ($H_2O$)
**Lane 10:** S41A8, *Set 8*, N.a R-23, Extr. 210 (formalin, muscle)
**Lane 11:** S42A8, *Set 8*, N.a R-8, Extr. 216 (formalin, muscle)
**Lane 12:** S43A8, *Set 8*, N.a R-8, Extr. 216 (formalin, muscle)
**Lane 13:** *Set 8*, N.a R-7, Ex. 224 (formalin, muscle)
**Lane 14:** *Set 8*, N.m R-5, Ex. 222 (formalin, muscle)
**Lane 15:** *Set 8*, N.m R-13, Ex. 223 (formalin, muscle)
**Lane 16:** *Set 8*, N.m R-5, Ex. 302 (formalin, vertebra with spine)
**Lane 17:** *Set 8*, "– " PCR, ($H_2O$)
**Lane 18:** *Set 6*, N.a R-7, Ex. 224 (formalin, muscle)
**Lane 19:** S44A6, *Set 6*, N.a R-23, Ex. 210 (formalin, muscle)

**Gel (b):**

**Lane 1:** S45A6, *Set 6*, N.a R-8, Ex. 216 (formalin, muscle)
**Lane 2:** *Set 6*, N.m R-5, Ex. 222 (formalin, muscle)
**Lane 3:** *Set 6*, N.m R-13, Ex. 223 (formalin, muscle)
**Lane 4:** *Set 6*, N.m R-5, Ex. 302 (formalin, vertebra with spine)
**Lane 5:** *Set 6*, "– " PCR, ($H_2O$)

**Ex.** – identifier number for DNA extract (e.g. Ex. 302 is extract no. 302)
**R** – identifier for fish individual (e.g., R-5 is fish no. 5)
**underlined** – PCR amplification sample sent for sequencing (e.g., S40MM is amplicon sample no. 40 (S40), which was generated with mitochondrial primer set 16S and with DNA of *Nezumia micronychodon* (N.m))

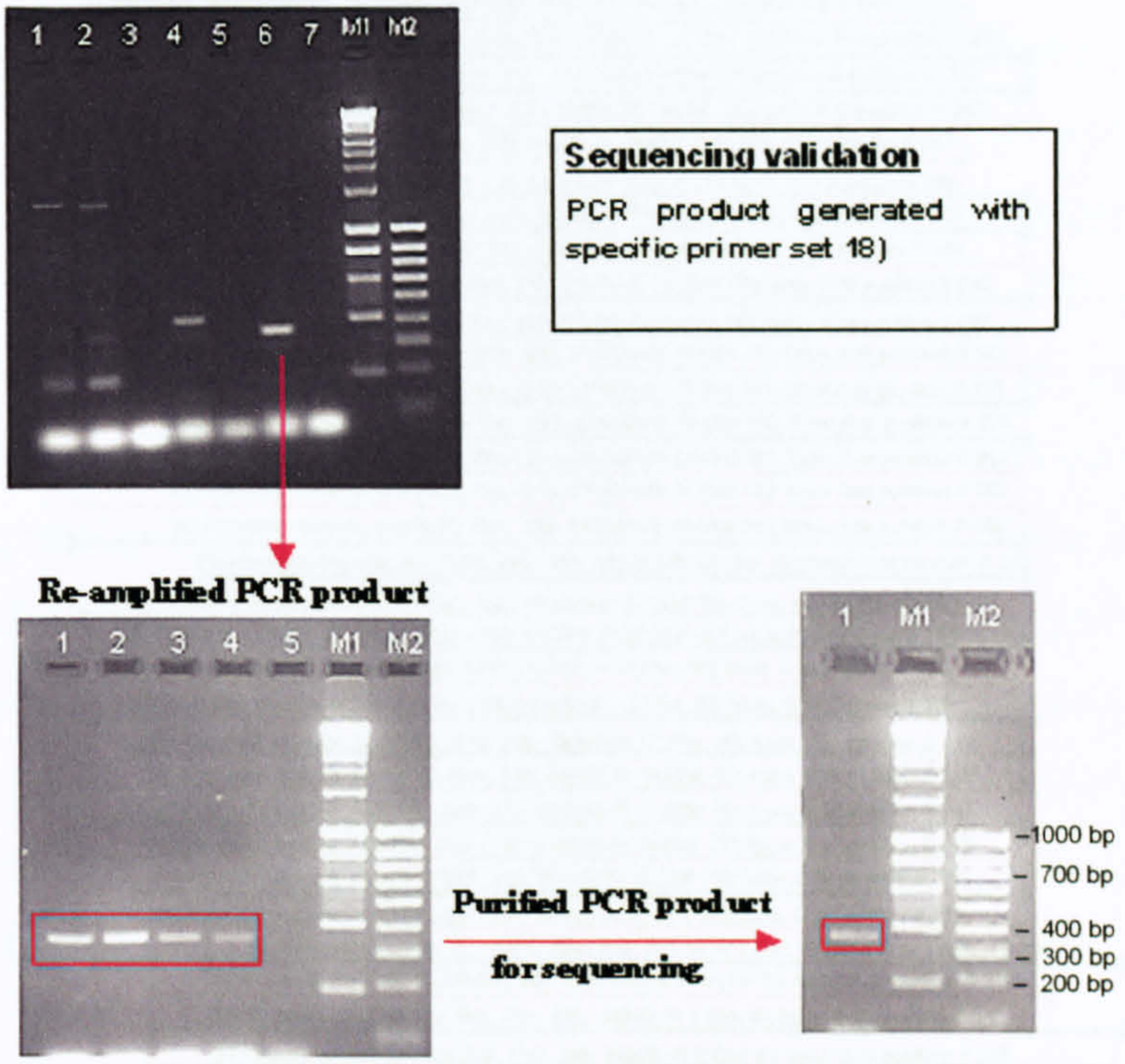



**Figure 43** – PCR product size of ~350 bp generated by the use of designed (RAPD-derived) *Nezumia*-specific primer set 18 and DNA of *Nezumia aequalis* extracted by phenol protocol (protocol (H)) from 70% ethanol preserved muscle tissue. Sequencing and sequence validation was carried out at the Natural History Museum, London.
"M1" indicates quantitative size marker fragments – HyperLadder I (Bioline)
"M2" indicates quantitative size marker fragments – HyperLadder IV (Bioline)

In this study, the reproducibility of sequencing results was confirmed by direct sequencing of PCR products generated from different formalin-arDNA extracts produced from different tissue types of one fish individual (these sequences were 100% identical), from different formalin-fixed, Steedman's-preserved fish individuals (specimens), and from differently preserved *Nezumia* specimens (DMSO [Fig. 41(c)-L: 8] and 70% ethanol preserved [Fig. 43]). Sequence alignment (individual sequence reads) is shown in Fig. 44. DNA extractions, PCR and sequencing experiments were carried out in different laboratories, so the possibility of cross-contamination by chemicals and DNA present in a particular laboratory was eliminated, and it was possible to validate them independently.

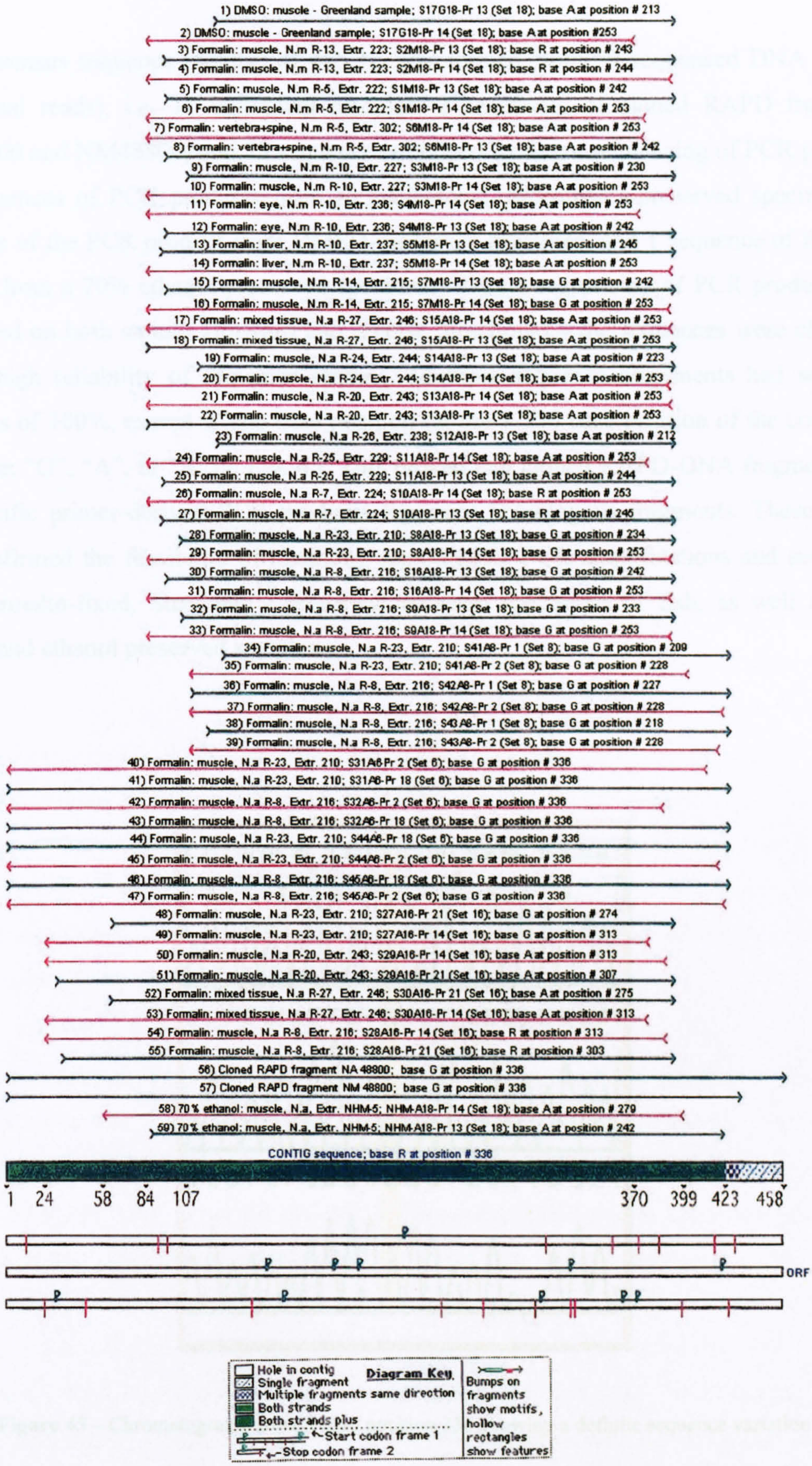



**Figure 44** – Multiple alignments of nucleic acid sequences (individual sequence reads) of cloned RAPD-PCR fragments and PCR products (including DMSO and ethanol validation specimens).

The consensus sequence (contig) of 485 bp was formed from 59 sequenced DNA strands (individual reads), i.e. 31 individual samples: 2 sequences of cloned RAPD fragments (NA48800 and NM48800) and 29 sequences obtained by direct sequencing of PCR products (27 sequences of PCR products from formalin-fixed, Steedman's-preserved specimens, 1 sequence of the PCR product from DMSO-preserved specimen, and 1 sequence of the PCR product from a 70% ethanol-preserved specimen). Direct sequencing of PCR products was performed on both strands (forward and reverse directions) – the sequences were clear and with a high reliability of sequence reading. Multiple sequence alignments had sequence identities of 100%, except at one base position (at the # 336 base position of the consensus sequence: "G", "A", or "R") – Fig. 45. This proved that cloned RAPD-DNA fragments and the specific primer-derived PCR products represent homologous fragments. These results also confirmed the feasibility of obtaining reproducible PCR amplifications and sequences from formalin-fixed, Steedman's preserved museum specimens of fish, as well as from DMSO and ethanol preserved samples.

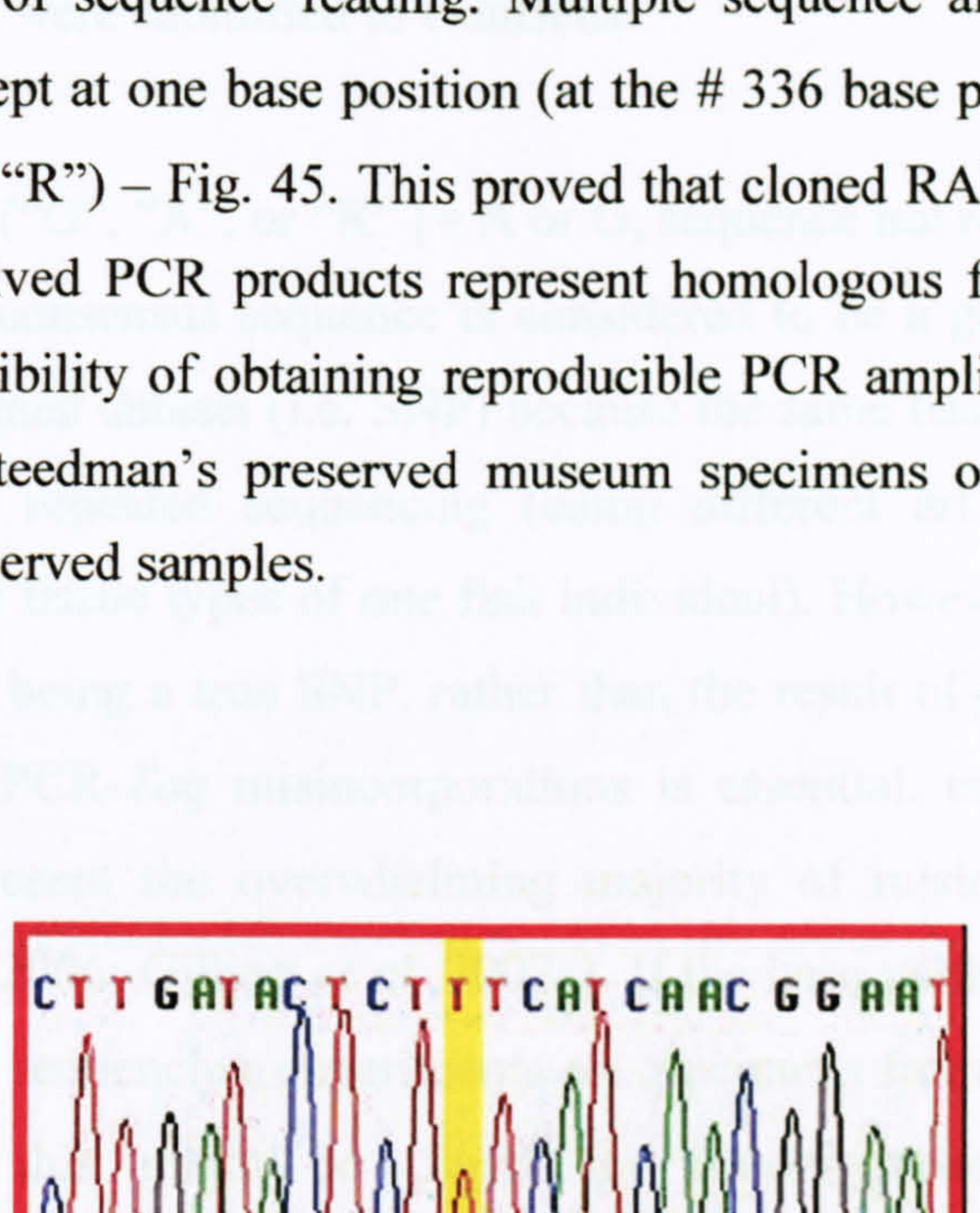



**Figure 45** – Chromatograms around base position 336 showing a definite sequence variation.

### 3.2.5. Sequence analysis and submission to GenBank

Selected DNA sequences of *Nezumia aequalis* and *N. micronychodon* were submitted to the GenBank database under accession numbers: AY826775 - AY826792 as reproducible sequences (Appendices 13 and 14). Only sequences with absolutely reliable reading, that is sequences obtained from arDNA extracts that were checked against primer sets that should not, and did not, generate PCR products (because primers were completely or partly in the sequence of the vector), were submitted to GenBank.

One base disagreement ("G", "A", or "R" [= A or G, sequence not readable]) at the 336 base position of the 458 bp consensus sequence is considered to be a genuine nucleic acid base variation across the aligned dataset (i.e. SNP) because the same base was observed from the one fish individual in repeated sequencing (using different arDNA extracts that were extracted from different tissue types of one fish individual). However, determination of this SNP (A/G, i.e. T/C) as being a true SNP, rather than the result of postmortem/preservation sequence damage and PCR-*Taq* misincorporations is essential, especially since C-T/G-A miscoding lesions represent the overwhelming majority of misincorporations in ancient DNA (see Stiller *et al.* 2006; Gilbert *et al.* 2007a). If the base variation is proved to be real by performing multiple sequencing experiments on specimens from other collections and/or fresh/frozen samples, this might be useful for investigations of single nucleotide polymorphism (SNP) in *Nezumia* species.

The same sequence is present in both investigated species suggesting that this section belongs to a well-conserved region of the genomes. The sequence might be a part of some gene responsible for important phenotypic features (e.g. production of enzyme, or other protein, important for organism function). This assumption might be supported by the presence of the open reading frame (ORF) in the 458 bp-long sequence - Fig. 44. The Sequencher protein translation of this ORF provided a 152 amino acid sequence that was used for a BLASTP search in the NCBI protein database (Appendix 15). The translated amino acid sequence (by Sequencher) aligned against the GenBank (NCBI) protein dataset (using the BLASTP alignment) revealed a relatively high probability (e-values: 8e-11 and 2e-10) that this ORF is a part of the conserved Rhomboid protein family that contains integral membrane proteins with strongly conserved histidines. Similar results were also

found with BLASTX search (nucleic acid/protein) against the GenBank (NCBI) protein database (Appendix 16).

The 485 bp long consensus (contig) DNA sequence was subjected to BLASTN search (nucleic acid/nucleic acid alignment; Fig. 46). The BlastN search with a 458 bp consensus DNA sequence against the GenBank (NCBI) nucleic acid database found no similar sequence, i.e. matching sequences only from this study – red lines (Fig. 46). The highest sequence identities with other organisms that were deposited in the GenBank nucleic acid database were found in a mouse BAC clone with 25 bases (100% similarity), in *Homo sapiens* with 22 bases, and zebrafish with 21 bases. These similarities occurred in different regions of the *Nezumia* sequence.

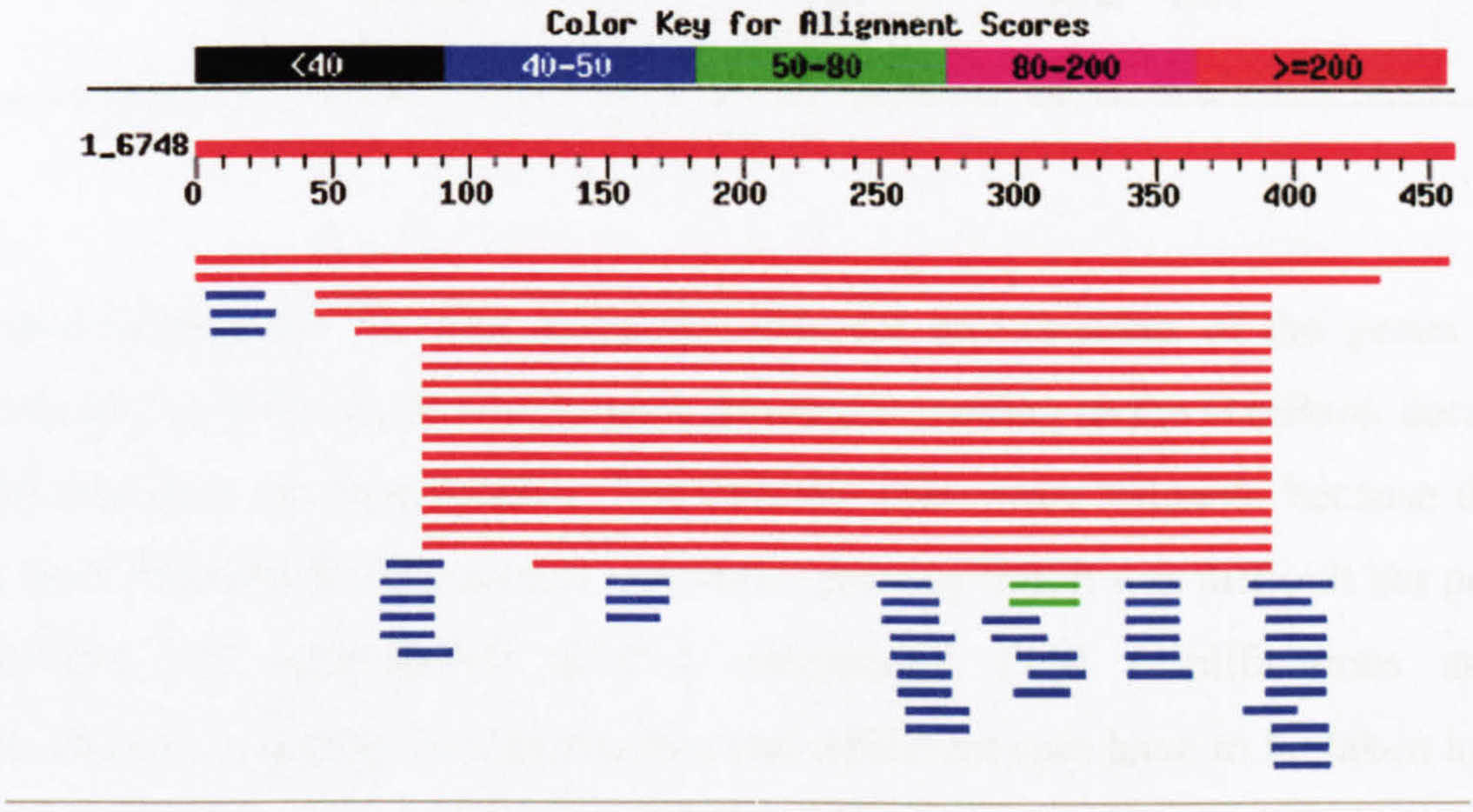
**Figure 46** – BlastN search with the 458 bp consensus DNA sequence against the GenBank (NCBI) nucleic acid database matches only sequences from this study (red lines).

***NUCLEIC ACID/ NUCLEIC ACID BLAST SEARCH:***

NCBI *results of* **BLAST**

**BLASTN**

**Query=** (458 letters)



Another revealed sequence (a 358 bp long DNA sequence of the NA48500B cloned RAPD fragment), which was not incorporated into the previously described consensus (contig) sequence of 458 bp, was also subjected to nucleic acid/nucleic acid BLAST alignments against the GenBank database (NCBI) – Fig. 47. A BLASTN search with this 358 bp long DNA sequence also revealed no significant matches, apart from itself (red line). This 358 bp long DNA sequence was also submitted to the GenBank database under accession number AY826774.

**Figure 47** – BlastN search with the 358 bp long DNA sequence of the NA48500B cloned RAPD fragment against the GenBank (NCBI) nucleic acid database reveals no significant matches

***NUCLEIC ACID/ NUCLEIC ACID BLAST SEARCH:***


These sequence data are the first molecular data for any member of the genus *Nezumia* (which currently contains approximately 60 species) according to the GenBank database and other published data on these species. The data are even more valuable because they were extracted from formalin-fixed museum specimens proving that it was difficult but possible to obtain reliable and reproducible arDNA extractions, PCR amplifications and DNA sequences. However, appropriate approaches and sufficient care have to be taken in order to obtain these kinds of data. A developed and adopted strategy for this study on *Nezumia* species is shown in Figure 48.

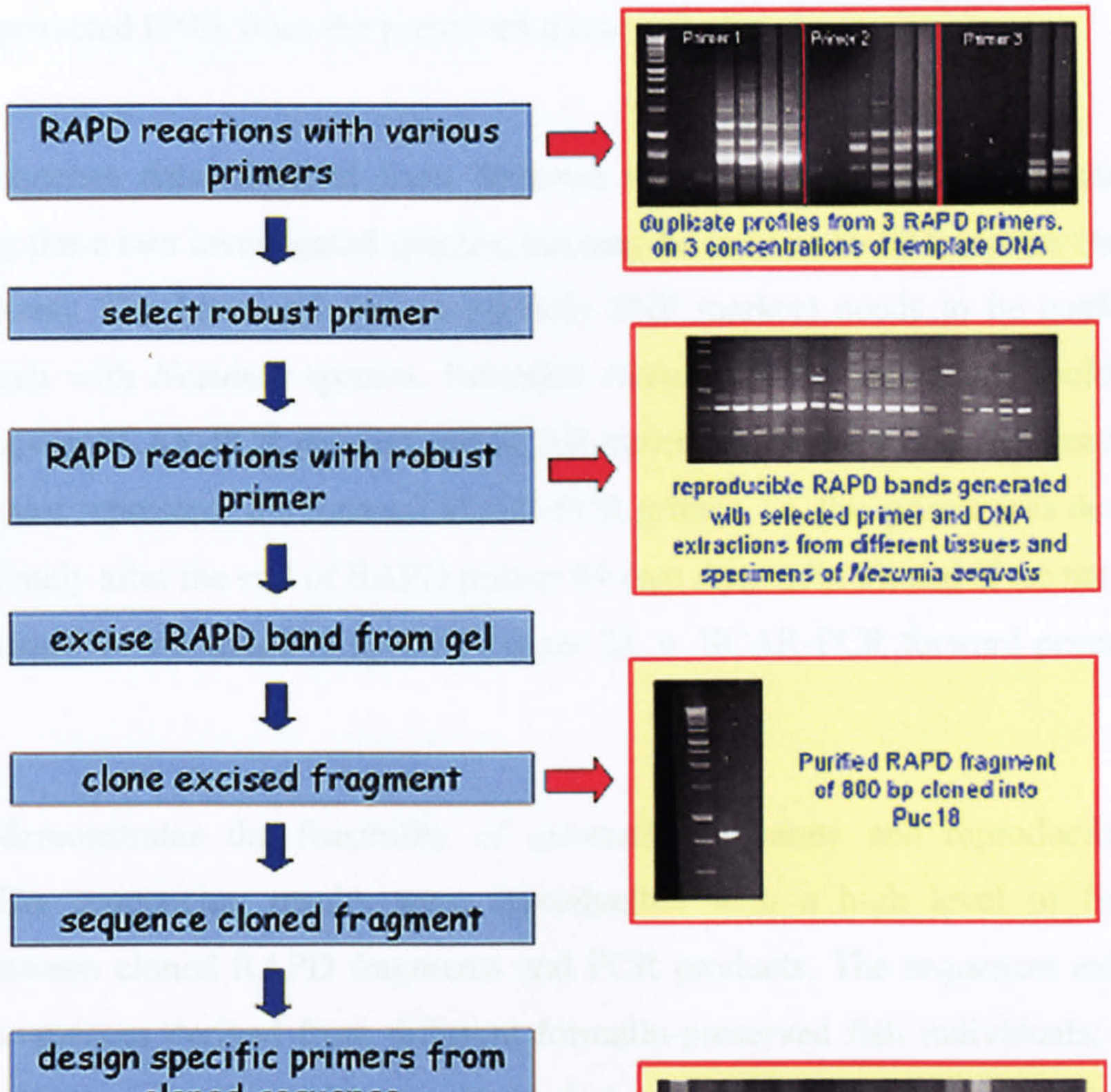



**Figure 48** – Adopted strategy for developing the specific RAPD-PCR derived primers used for generating the specific (anonymous) targeted sequences that were subjected to direct sequencing.

Generally, this kind of molecular investigation is difficult to perform on formalin-fixed, Steedman's-preserved fish specimens, especially on species with no available molecular information. The main reason is the prolonged damaging effect of formalin on the quality and yield of extracted DNA from the preserved tissue samples.

The DNA sequence data obtained from *Nezumia* do not present molecular markers for distinguishing these two investigated species, but may be used as an STS marker (sequence-specific markers). The base variation (a possible SNP marker) needs to be confirmed in further research with *Nezumia* species. Revealed *Nezumia* DNA sequences could also be used for designing SCAR-PCR primers and SCAR molecular markers. The designed specific primer 21 almost represents the forward SCAR-PCR primer, i.e. this primer was designed to begin immediately after the end of RAPD primer 44 (see Appendix 9), and if the ten bases of the RAPD primer 44 are added to specific primer 21, a SCAR-PCR forward primer would be created.

This study demonstrates the feasibility of generating accurate and reproducible DNA sequences. The sequencing results were reproducible with a high level of fidelity of sequences between cloned RAPD fragments and PCR products. The sequences exhibited a match with sequences derived from different formalin-preserved fish individuals, different arDNA extractions, as well as with sequences derived from DMSO and ethanol-preserved fish specimens of *Nezumia*. All sequences were identical, except at a single base position (A/G) at exactly at the same location in the sequences. The probability that an artificial mutation and/or misincorporation happened always at the same base position in the sequence (after repeated sequencing of the same and different arDNA extracts [DNA extracted from different tissue types] from the one fish individual and by using different PCR primers for amplifying the same genome region) is low, unless hypotheses of "hotspot damage" in aDNA (Gilbert *et al.* 2005b, Binladen *et al.* 2006) and non-random induced DNA damage by different factors (heating - Banerjee and Brown (2004); post-PCR handling and sequencing chemistry - Brandstatter *et al.* (2005)) are also true for preserved specimens (see Tang 2006; Wandeler *et al.* 2007). These require further research with experiments designed specifically to investigate these issues.

Contradictory findings related to the accuracy and fidelity of sequences derived from formalin-fixed tissue is reported by many researchers. For example, accurate and reliable

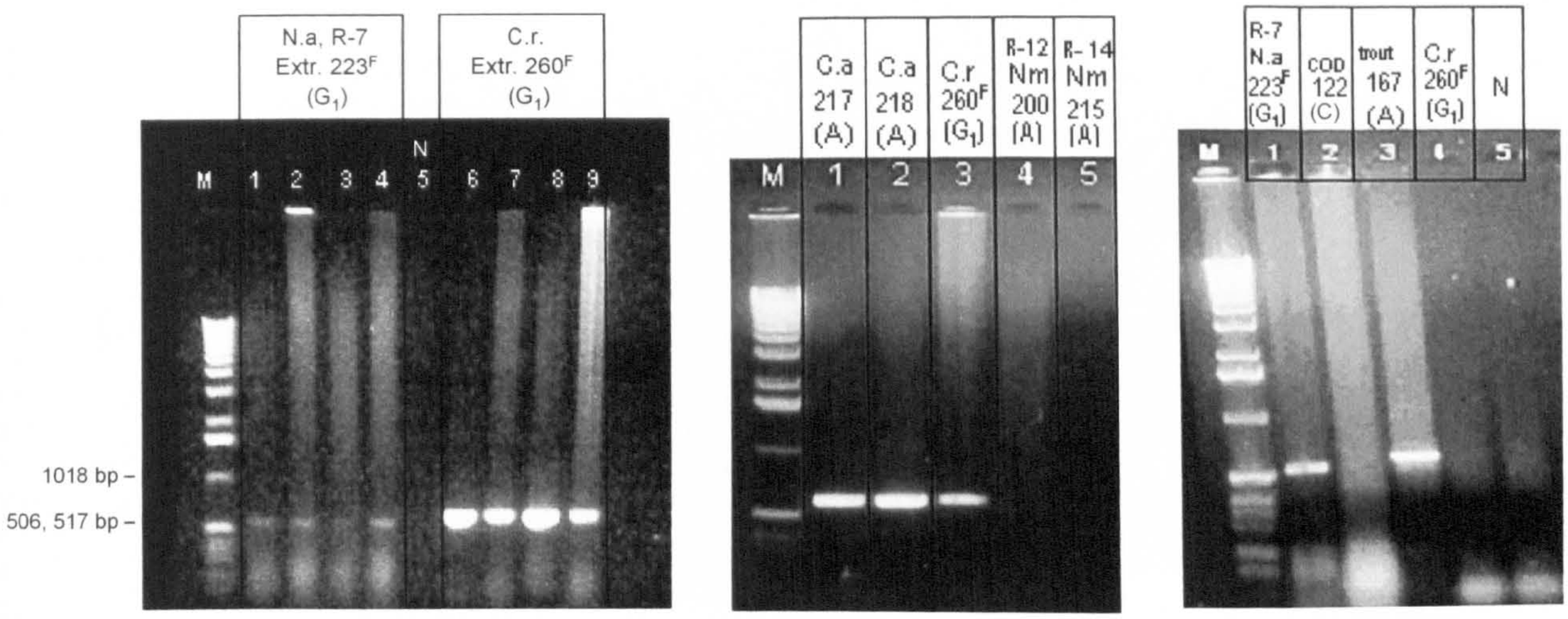
DNA sequence information was obtained in studies of Dubeau *et al.* (1986), France and Kocher (1996), Shedlock *et al.* (1997), Chase *et al.* (1998a), Quattro *et al.* (2001), Boyle *et al.* (2004), Zardus *et al.* (2006), whereas artifactual base mutations (located throughout the sequence without a particular "hot spot") have been reported by De Giorgi *et al.* (1994), Williams *et al.* (1999), Quach *et al.* (2004). The issue about the reliability of the sequence information obtained from formalin-fixed samples is still the subject of debates among researchers (see Nadler (1999) and Littlewood (1999) published correspondence to emphasise the problem; Tang (2006) workshop summary report on formalin-fixed biological samples).

A consistency and fidelity of sequencing results in this study is related to the 458 bp-long, well-conserved and anonymous genome region of *Nezumia*, but this does not necessarily mean that all genome regions would behave in the same way. This might be related to a specific base composition, or specific position of this region in a genome that is less susceptible to formalin-induced alternations of the sequence, or to the use of particular thermostable DNA polymerases for PCR experiments.

## 3.3. Mitochondrial DNA

Investigations were carried out on different regions of the mitochondrial genome (351 bp region of the cyt b, 605 bp of COII, ~500 bp of COIII, 433 bp of 12S, 146 bp of COI+ tRNA+COII, and 570 bp of 16S) on formalin-fixed, Steedman's-preserved museum specimens of *Nezumia*. Additional evidence on control specimens of DMSO- and ethanol-preserved samples of *Nezumia*, ethanol preserved *Coryphaenoides*, fresh/frozen samples of cod and rainbow trout, and 10 differently preserved specimens of mackerel is also presented in this section.

An unexpected problem during this project was the extreme difficulty in amplifying mitochondrial DNA sequences from **formalin-arDNA of *Nezumia*** (see Appendix 17). A clear example of this difficulty is a PCR experiment shown in Fig. 49(a) - 16S mt-PCR amplifications with formalin-arDNA of *Nezumia aequalis* (Lanes: 1-4) and ethanol-arDNA of *Coryphaenoides rupestris* (Lanes 6-9). PCRs were performed under the same conditions (including the use of the same DNA extraction protocol ($G_I$) for producing these two arDNA extracts), but the results obtained were very different. PCR products of expected size (~ 570 bp) were hardly visible on the gel with formalin-arDNA of *Nezumia aequalis* (Fig. 49, Lanes: 1-4). Increased Mg and/or PCR-DNA template concentrations did not improve PCR yield. Nor did an addition of Q solution (Qiagen PCR kit) make significant improvements in PCR performance. In contrast to these PCR results with formalin-arDNA of *Nezumia*, the 16S mt-PCR amplifications generated PCR products of a very good yield with DNA of ethanol preserved *Coryphaenoides rupestris* (Fig. 49(a), Lanes 6-9).



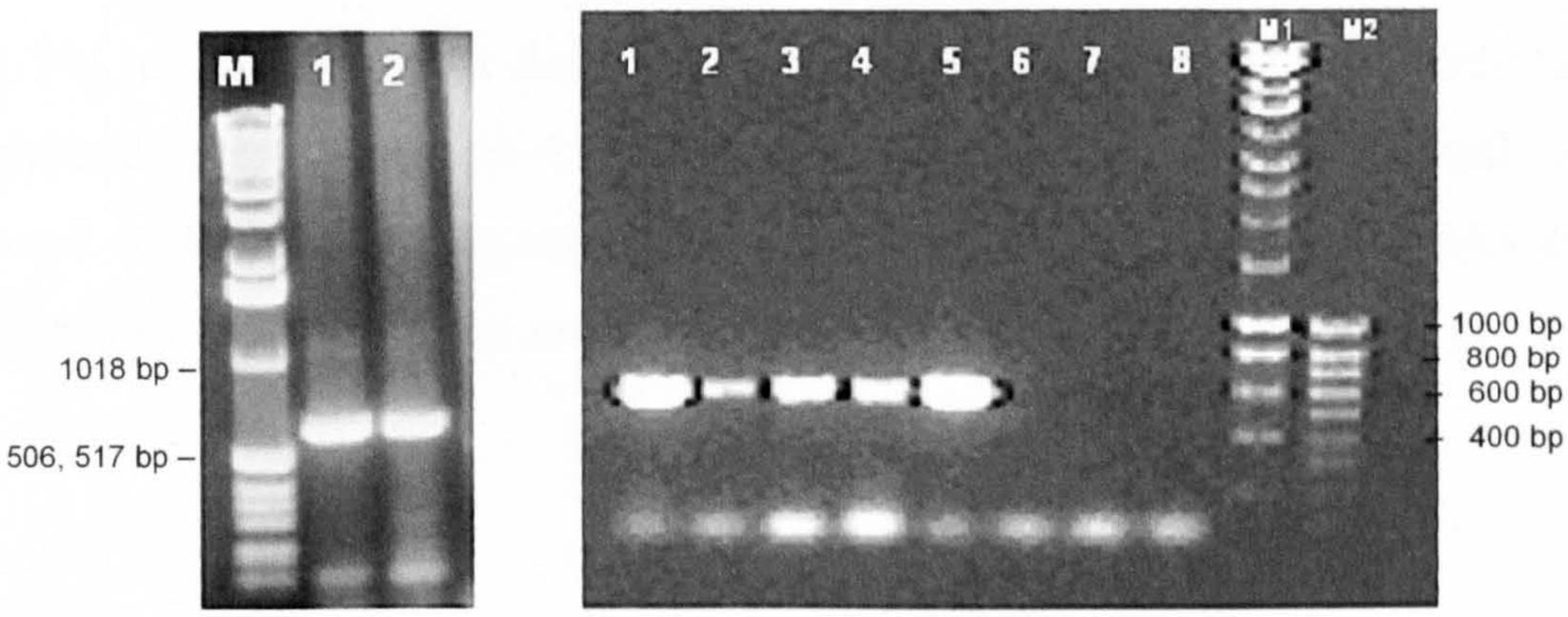

**(a) - Southampton** **(b) - Nescot** **(c) - Nescot**

**Figure 49** – PCR amplifications of the 16S mitochondrial gene (region of ~570 bp) with formalin-arDNA of *Nezumia aequalis* (N.a), formalin-arDNA of *Nezumia micronychodon* (N.m), ethanol-arDNA of *Coryphaenoides rupestris* (C.r), ethanol-arDNA of *Coryphaenoides armatus* (C.a), and DNA from fresh/frozen sample of cod and trout. PCRs were performed in two different laboratories, using different extraction protocols, *Taq* kits from different suppliers, PCR conditions and PCR thermal cyclers.
"M" indicates size marker fragments - 1 Kb DNA Ladder (Gibco).

These experiments were performed in the Southampton Laboratory using Ampli*Taq* DNA polymerase (Perkin Elmer) and Perkin Elmer 480 thermal cycler (annealing temperature 45°C for 1 min). Using the same arDNA extracts of *Nezumia aequalis* (no. $223^F$) and *Coryphaenoides rupestris* (no. $260^F$) and performing PCRs under different conditions and in a different molecular laboratory (the Nescot Laboratory - the main laboratory for the study), i.e. using different *Taq* polymerase (Hybaid- AGS Gold*Taq*) and a different thermal cycler (Hybaid Omn-E with hot lid; annealing temperature 40°C for 1 min), 16S PCR products were also obtained (Fig. 49(b)-Lane: 3 for *Coryphaenoides* and Fig. 49(c)-Lane: 1 for *Nezumia*). If a 100x diluted suspension of *Coryphaenoides* ethanol-arDNA extract no. $260^F$ was applied, the PCR product was absent (Fig. 49(c)-Lane: 4) probably because of a too low PCR-DNA template concentration. Ethanol-arDNA extracts (nos. 217 and 218) of *Coryphaenoides armatus* produced by protocol (A) also gave successful 16S mt-PCR amplifications (Fig. 49(b)-Lanes: 1 and 2). Unsuccessful 16S mt-PCR with control DNA of cod (Fig. 49(c)-Lane: 2) might be related to 72 hours of washing the frozen tissue sample in 1xGTE buffer prior to DNA extraction and application of phenol protocol (C), or too much DNA in the PCR reaction (over 100 ng DNA) inhibiting the PCR.

Considering the suitability of a particular DNA extraction protocol to a specific molecular marker system, it is important to mention that formalin-arDNA extract no. $223^{F}$ gave successful 16S mt-PCR amplifications, but not successful RAPDs (see Fig. 19 - Lane: 14 in gels (a) and (b); p 130). Formalin-arDNA extract no. 215 (produced by protocol (A)) did not generate 16S mt-PCR product (Fig. 49(b)-Lanes: 5), but did generate RAPDs (Fig. 22 (b); p 145). However, ethanol-arDNA extracts of *Coryphaenoides* nos. 217 and 218, produced by protocol (A), generated successful 16S mt-PCRs (Fig. 49(b)-Lanes: 1 and 2) and RAPDs as well (Fig. 17; p 126). This might suggest that selection of a DNA extraction protocol for short-term ethanol-preserved specimens is not so important for successful PCR amplifications of different markers, but it seems that an appropriate selection of a DNA extraction protocol for amplifying particular markers is of crucial importance for formalin-Steedman's preserved specimens.

If 16S PCR amplifications were performed with **arDNA extracts from DMSO**- (Fig. 50(a)) and **ethanol-preserved samples of *Nezumia*** (Fig. 50(b)), PCR products of expected size (~570 bp) were generated without difficulty. The same DMSO-arDNA extracts (nos. 300 and 301), produced by protocol (A), generated RAPDs and PCR products with *Nezumia*-specific primer set 18 (Fig. 16; p 126).




**(a) - DMSO** **(b) - ethanol-arDNA of *Nezumia***

**Figure 50** – The 16S mitochondrial PCR amplifications with **(a)** DMSO-arDNA of *Nezumia* cf. *aequalis*, and **(b)** ethanol-arDNA of *Nezumia aequalis* (L: 1, 3, 5) and *N. micronychodon* (L: 2, 4).
"M" indicates size marker fragments - 1 Kb DNA Ladder (Gibco).
"M1" indicates quantitative size marker fragments – "Bioline" HyperLadder I
"M2" indicates quantitative size marker fragments – "Bioline" HyperLadder IV

Generally, the primer sets for six tested mitochondrial genes barely gave any results on formalin/Steedman's-arDNA of *Nezumia* (Fig. 51; Appendix 17). According to the results the mitochondrial primer **Set *1 (cyt b)*** – (Fig. 51 (a)-L: 1-5, 6), and possibly, **sets: *2 (CO II)*** – (Fig. 51 (a)-L: 7-10, 12), ***4 (12S)*** – (Fig. 51 (a)-L: 18-19, (b)-L: 1-2, 4) and ***5 (COI + tRNA +COII)*** – (Fig. 51 (b)-L: 5-8, 10) were not the best choice of primers for macrourid fishes. However, the mitochondrial primer **Sets: *3 (CO III)*** – (Fig. 51 (a)-L: 13-17, (c)-L: 1-5, 7) and ***16S*** (Fig. 49, 50, 51(e)) were definitely the right choices of mitochondrial primers (regarding a primer match) for macrourid and other fishes, but the results obtained with formalin-arDNA of *Nezumia* were very poor.

Using mt-Set 3 (COIII), only 7% of successful PCR amplifications were achieved with formalin-arDNA of *Nezumia micronychodon* (4 out of 61 tested PCRs) and 2% with formalin-arDNA of *Nezumia aequalis* (1 out of 48 tested PCRs). With ethanol-arDNA of *Coryphaenoides armatus*, the success of PCR amplifications with the mt-Set 3 was 100% (10 out of 10 tested).

Applying the mt-Set 16S, 18% successful PCR amplifications were produced with formalin-arDNA of *Nezumia micronychodon* (9 out of 50 tested PCRs) and 30% with formalin-arDNA of *Nezumia aequalis* (18 out of 60 tested PCRs). The success of 16S mt-PCR amplifications with ethanol-arDNA of *Nezumia micronychodon* and *Nezumia aequalis* was 67% - 100% (it would be 100%, but PCRs with the problematic DNA extract no. NHM-2 [see Fig. 20; p 136] and PCRs that were performed with too high a concentration of PCR-DNA template [over 100 ng DNA in PCR reaction] were included in the final calculations). The success of PCR amplifications with ethanol-arDNA of *Coryphaenoides armatus* and *Coryphaenoides rupestris* using mt-Set 16S was 83% - 100%.

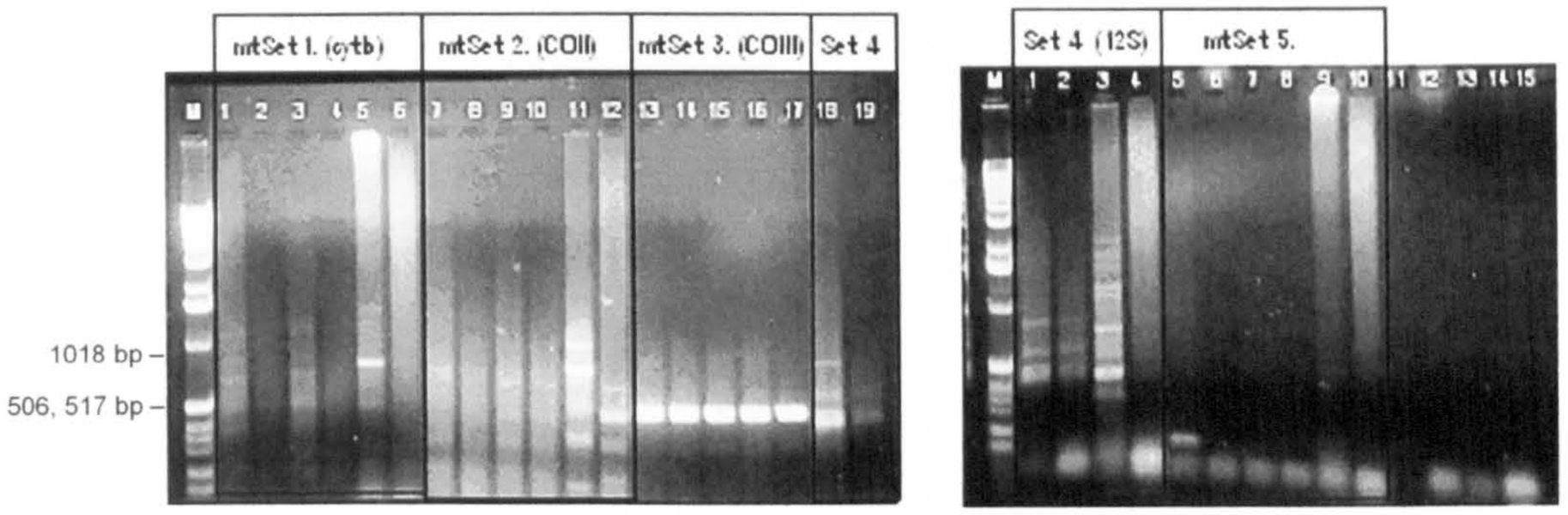



**Figure 51** – The mitochondrial PCR amplifications with the six tested mitochondrial genes:

(a) **Lanes 1-4:** ethanol-arDNA of *Coryphaenoides armatus* with mt-Set 1. (cyt b)
**Lane 5:** DNA of frozen cod with mt-Set 1. (cyt b)
**Lane 6:** formalin-arDNA of *N. micronychodon* with mt-Set 1. (cyt b)
**Lanes 7-10:** ethanol-arDNA of *Coryphaenoides armatus* with mt-Set 2. (COII)
**Lane 11:** DNA of frozen cod with mt-Set 2. (COII)
**Lane 12:** formalin-arDNA of *N. micronychodon* with mt-Set 2. (COII)
**Lanes 13-17:** ethanol-arDNA of *Coryphaenoides armatus* with mt-Set 3. (COIII)
**Lanes 18-19:** ethanol-arDNA of *Coryphaenoides armatus* with mt-Set 4. (12S)

(b) **Lanes 1-2:** ethanol-arDNA of *Coryphaenoides armatus* with mt-Set 4. (12S)
**Lane 3:** DNA of frozen cod with mt-Set 4. (12S)
**Lane 4:** formalin-arDNA of *N. micronychodon* with mt-Set 4. (12S)
**Lanes 5-8:** ethanol-arDNA of *Coryphaenoides armatus* with mt-Set 5. (COI+ tRNA+ COII)
**Lane 9:** DNA of frozen cod with mt-Set 5. (COI+ tRNA+COII)
**Lane 10:** formalin-arDNA of *N. micronychodon* with mt-Set 5. (COI+ tRNA+COII)
**Lanes 11-15:** negative PCRs with tested mt-primer Sets (sd$H_2O$ used as a PCR template)

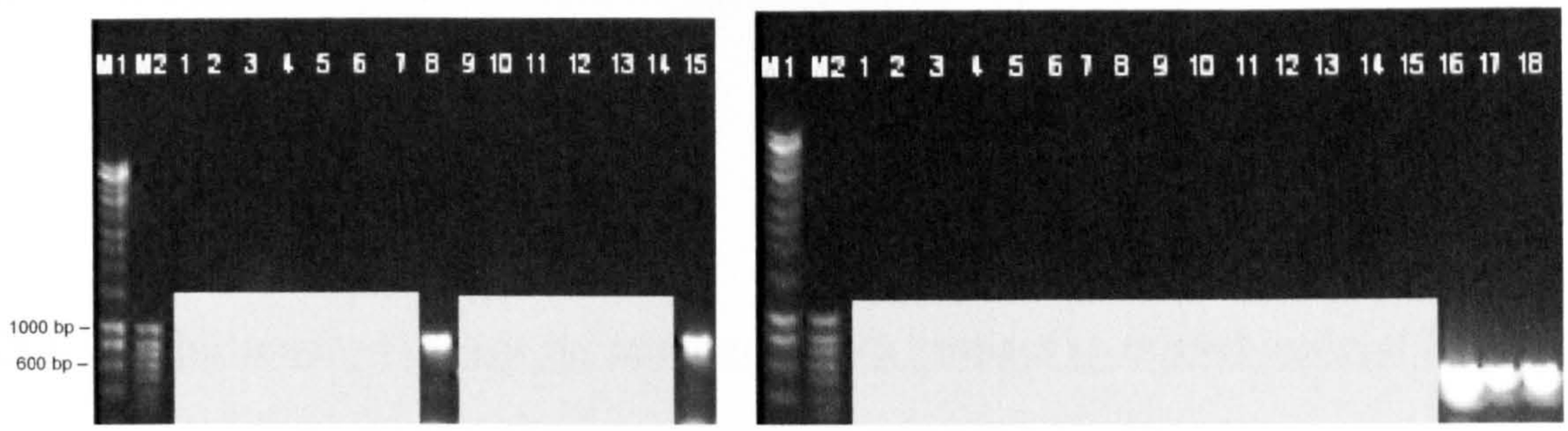
(c) **Lanes 1-5:** ethanol-arDNA of *Coryphaenoides armatus* with mt-Set 3. (COIII)
**Lane 6:** DNA of frozen cod with mt-Set 3. (COIII)
**Lane 7:** formalin-arDNA of *N. micronychodon* with mt-Set 3. (COIII)
**Lanes 8:** negative PCRs with tested mt-primer Sets (sd$H_2O$ used as a PCR template)

(d) DNA of fresh/frozen rainbow trout with: *mt-Set 1.* (**Lanes 2-3**), *mt-Set 2.* (**Lanes 4-5**), *mt-Set 3.* (**Lanes 6-7**), *mt-Set 4.* (**Lanes 8-9**), *mt-Set 5.* (**Lanes 10-11**).

(e) **Lane 1-10:** formalin-arDNA of *N. micronychodon* with the mitochondrial primer *Set 16S*
**Lane 11-17:** formalin-arDNA of *N. aequalis* with the mitochondrial primer *Set 16S*

During validation (in the Molecular Laboratory of the Natural History Museum London), another set of mitochondrial primers was tested for the cytochrome b gene region (the primers developed by Finnerty and Block (1992) for blue marlin) on **ethanol-preserved samples of *Nezumia***. This primer set produced good PCR products with ethanol-arDNA of *Nezumia* (Fig. 52 (a)). It will be worthwhile to test this primer set on formalin-arDNA of *Nezumia*.

**cytochrome b** (Finnerty and Block 1992)




**(a) – ethanol-arDNA of *Nezumia*** **(b) – mackerel (frozen tissue)**

**Figure 52** – The cytochrome b mitochondrial PCR amplifications (primers developed by Finnerty and Block 1992) with:

**(a)** ethanol-arDNA of *Nezumia* cf. *aequalis* **(Lane 8)** and ethanol-arDNA of *Nezumia* cf. *micronychodon* **(Lane 15)**. Both DNA extracts were produced by phenol protocol (H). Size of PCR product was ~800 bp.

**(b)** DNA extracted from fresh/frozen mackerel (*Scomber scombrus*) using the protocol (F) – Promega Kit (**Lanes: 16** (10x diluted DNA extract), **17** (50x diluted DNA extract), **18** (100x diluted DNA extract). Size of PCR product was smaller than in *Nezumia* (~600 bp).

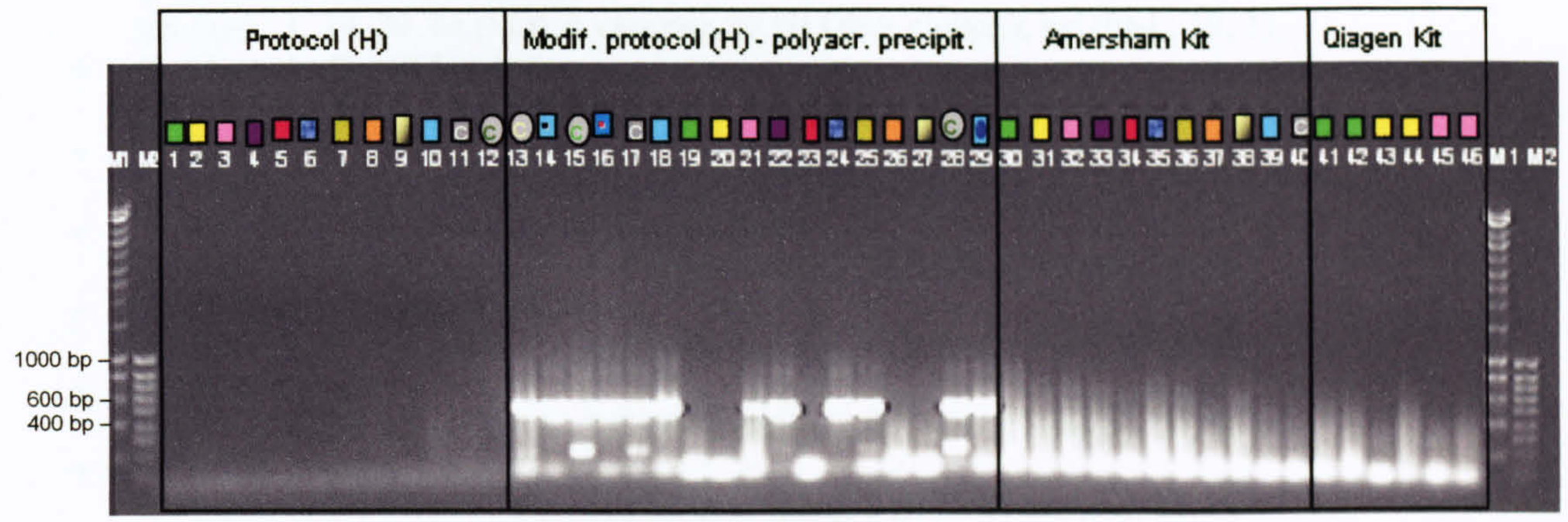
"M1" indicates quantitative size marker fragments – "Bioline" HyperLadder I
"M2" indicates quantitative size marker fragments – "Bioline" HyperLadder IV

Such poor results with formalin/Steedman-arDNA were unexpected because it is known that mitochondrial sequences are present in a high copy number in each cell and it was presumed that multiple copies of mtDNA would provide sufficient DNA for successful PCR and molecular investigations on *Nezumia*. At the moment, there is no clear explanation of why there is such great difficulty in amplifying mitochondrial DNA sequences of formalin-fixed, Steedman's-preserved specimens of fish, but results obtained in this study unquestionably indicate it. The only explanation could be some kind of cross-linking caused by formalin, and may be in interaction with other compounds in preservation solution, that did not allow the release of DNA from mitochondria, or caused severe damage of the mitochondrial genome of *Nezumia*.

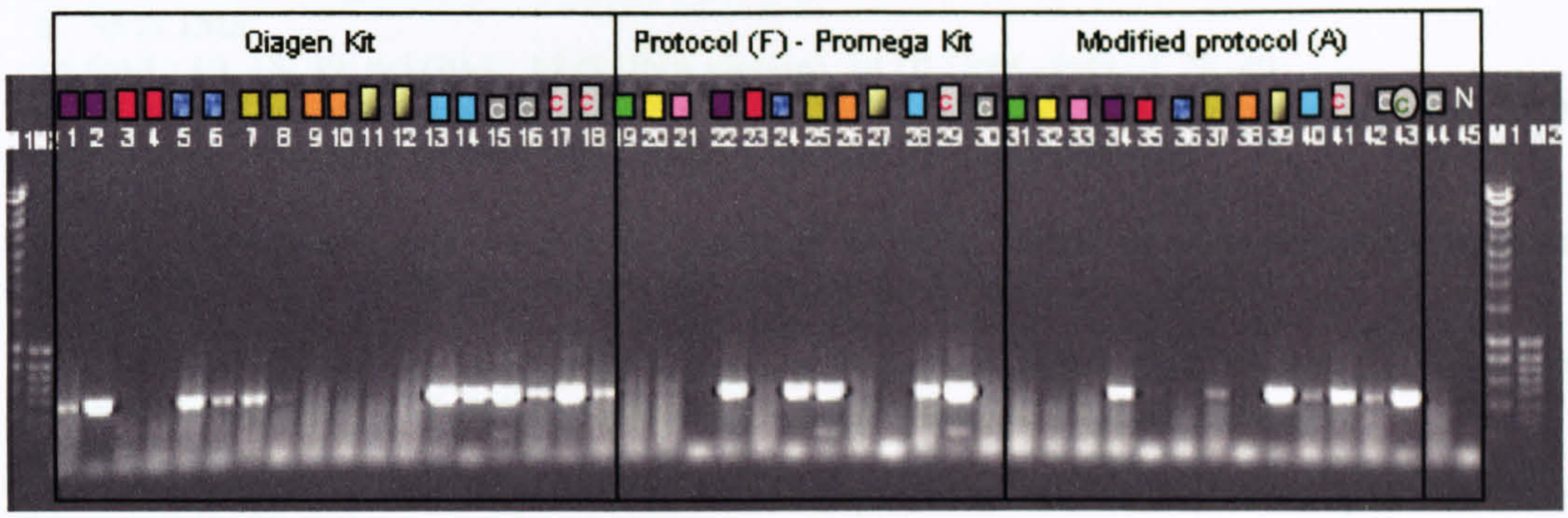
It was also found that the particular DNA extraction protocols (modified phenol protocols, especially a protocol ($G_1$)) might produce formalin-arDNA extracts that are more suitable for PCR amplifications of mitochondrial sequences (see the support material on CD), but more evidence is needed before drawing a definite conclusion. Protocol (A) developed in this project produced a few arDNA extracts that were suitable for mitochondrial PCR, but some modification might make this protocol more suitable for amplification of mitochondrial genes. Additional evidence on mackerel is presented to support this assumption about the relevance of a DNA extraction protocol, preservation method, and the success of PCRs with particular mitochondrial genes and genome regions.

### 3.3.1. Additional evidence on mackerel with respect to mitochondrial PCRs

In this section, mt-PCR experiments with the cytochrome b and 16S regions of mackerel are presented (DNA obtained from fresh/frozen and ten differently preserved museum specimens of this marine fish) – Fig. 53 and 54. DNA extraction protocols tested in these experiments with mackerel undoubtedly indicated its importance in extracting PCR amplifiable mt-arDNA. The smallest modification of a protocol could have a significant effect on the success of PCR amplifications and/or the possibility to amplify a particular mitochondrial region (marker). For instance, if protocol (H) was used with 3M NaAc (sodium acetate) precipitation of DNA (Fig. 53(a)-Lanes: 1-12), PCR products were absent even with fresh/frozen and ethanol preserved tissue samples. However, if the same phenol DNA extraction protocol was applied with polyacrylamide DNA precipitation instead of using 3M NaAc, the results were very different (Fig. 53(a)-Lanes: 13-29): most PCR amplifications yielded PCR product. This might be caused by some specific problem of this laboratory-made 3M NaAc solution, but nevertheless it indicates the importance of each step in a particular DNA extraction protocol. The same solution of 3M NaAc was used in extracting ethanol-preserved specimens of *Nezumia* – one of the arDNA extracts exhibited PCR inhibition, but four others were PCR amplifiable (although for three of these extracts, the DNA precipitation time was significantly reduced – from overnight to 45 min).




**(a)**




**(b)**

**Figure 53** – The cytochrome b mitochondrial PCR amplifications (primers developed by Finnerty and Block, 1992) with DNA of mackerel (*Scomber scombrus*) extracted by different DNA extraction protocols, using fresh/frozen and differently preserved muscle tissue samples. Muscle tissue was subjected to different washing/drying regimes prior to DNA extraction:

**A** - tissue was washed overnight in 1xTE buffer at room temperature prior to DNA extraction;

**B** - tissue was washed for a few minutes in $sdH_2O$, and then dried for a few minutes at room temperature prior to DNA extraction;

**C** - tissue was washed twice for 5 minutes in PBS (phosphate-buffered saline) and twice in $sdH_2O$ for a couple of minutes, and then briefly dried for a few minutes at room temperature prior DNA extraction

**D** - tissue was washed in 1xGTE buffer for 24 hours (gently rotating the samples) at room temperature and frequently changing the buffer, then washed in $sdH_2O$ for 10-15 minutes, and then briefly dried for a few minutes at room temperature prior to DNA extraction

**gel (a):**

**Lanes 1-12:** ***Protocol (H) - 3M NaAcetate (NaAc) DNA precipitation with applied tissue washing/drying regime A (indicated in the text below if different);***

**Lanes 13-29:** ***Modified protocol (H) - polyacrylamide DNA precipitation with applied tissue washing/drying regime A (indicated in the text below if different);***

**Lanes 30-40:** ***Amersham Bioscience Kit with applied tissue washing/drying regime B (indicated in the text below if different);***

**Lanes 41-46:** ***Qiagen Kit (tested both DNA elutions) with applied tissue washing/drying regime C (indicated in the text below if different);***

**gel (b):**

**Lanes 1-18:** ***Qiagen Kit (tested both DNA elutions) with applied tissue washing/drying regime C (indicated in the text below if different);***

**Lanes 19-30:** ***Promega Kit – Protocol (F) with applied tissue washing/drying regime B (indicated in the text below if different);***

**Lanes 31-43:** ***Modified protocol (A) – it was used liquid nitrogen instead of dry ice, applied tissue washing/drying regime D (indicated in the text below if different);***

**Lane 45:** ***PCR negative ($sdH_2O$ was used as PCR template)***

- **10% unbuffered formalin**
  gel (a)-L: 1, 19, 30, 41 (I DNA elution), 42 (II DNA elution); gel (b)-L: 19, 31
- **3% unbuffered formalin**
  gel (a)-L: 2, 20, 31, 43 (I DNA elution), 44 (II DNA elution); gel (b)-L: 20, 32
- **50% isopropanol**
  gel (a)-L: 3, 21, 32, 45 (I DNA elution), 46 (II DNA elution); gel (b)-L: 21, 33
- **70% IMS – industrial methylated spirits**
  gel (a)-L: 4, 22, 33; gel (b)-L: 1 (I DNA elution), 2 (II DNA elution), 22, 34
- **Modified Steedman`s solution**
  gel (a)-L: 5, 23, 34; gel (b)-L: 3 (I DNA elution), 4 (II DNA elution), 23, 35
- **70% ethanol**
  gel (a)-L: 6, 24, 35; gel (b)-L: 5 (I DNA elution), 6 (II DNA elution), 24, 36
- **100% ethanol**
  gel (a)-L: 7, 25, 36; gel (b)-L: 7 (I DNA elution), 8 (II DNA elution), 25, 37
- **3% buffered formalin – with 4% sodium tetraborate**
  gel (a)-L: 8, 26, 37; gel (b)-L: 9 (I DNA elution), 10 (II DNA elution), 26, 38
- **10% buffered formalin – with 4% sodium tetraborate**
  gel (a)-L: 9, 27, 38; gel (b)-L: 11 (I DNA elution), 12 (II DNA elution), 27, 39
- **95% IMS**
  gel (a)-L: 10, 18, 39; gel (b)-L: 13 (I DNA elution), 14 (II DNA elution), 28, 40
- **95% IMS (tissue was washed for 5 hours in 1xTE buffer at room temperature prior to DNA extraction) – *Modified protocol (H) - polyacrylamide DNA precipitation*;**
  gel (a)-L: 14,
- **95% IMS (tissue was washed for 1 minute in 1xTE buffer at room temperature prior to DNA extraction) – *Modified protocol (H) - polyacrylamide DNA precipitation*;**
  gel (a)-L: 16,
- **95% IMS (tissue was only dried at room temperature without any washing in 1xTE buffer or in other solution) – *Modified protocol (H) - polyacrylamide DNA precipitation*;**
  gel (a)-L: 16,
- **95% IMS (tissue was only dried at room temperature without any washing in 1xTE buffer or in other solution) – *Modified protocol (H) - polyacrylamide DNA precipitation*;**
  gel (a)-L: 29,
- **fresh/frozen**
  gel (a)-L: 11,17, 40; gel (b)-L: 15 (I DNA elution), 16 (II DNA elution), 30 (10x diluted), 42, 44 (Promega Kit; 50x diluted DNA extract),
- **fresh/frozen (tissue was washed for 5 hours in 1xTE buffer at room temperature prior to DNA extraction) – *Modified protocol (H) - polyacrylamide DNA precipitation*;**
  gel (a)-L: 13,
- **fresh/frozen (tissue was washed for 1 minute in 1xTE buffer at room temperature prior to DNA extraction) – *Modified protocol (H) - polyacrylamide DNA precipitation*;**
  gel (a)-L: 15,
- **fresh/frozen (tissue was not washed at all, i.e. the tissue was used for DNA extraction straight from the freezer) – *Unmodified and modified protocol (H) - 3M NaAc and polyacrylamide DNA precipitation*;**
  gel (a)-L: 12 (protocol (H) with 3M NaAc), 28 (protocol (H) with polyacrylamide); gel (b)-L: 43 (modified protocol (A)),
- **control 100% ethanol preserved – preserved in ethanol only for a few days**
  gel (b)-L: 17 (I DNA elution), 18 (II DNA elution), 29, 41

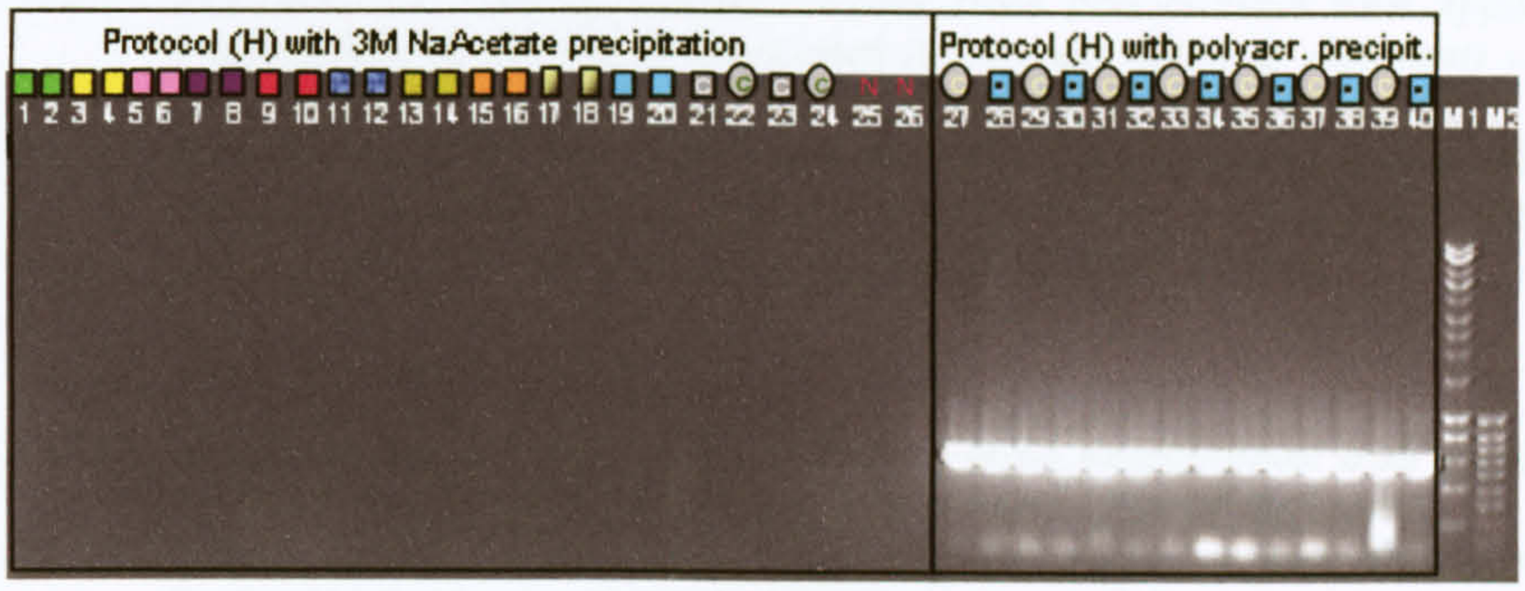
"M1" indicates quantitative size marker fragments – "Bioline" HyperLadder I
"M2" indicates quantitative size marker fragments – "Bioline" HyperLadder IV

Generally, different protocols produced arDNA that had different successes in amplifying the cytochrome b region of the mitochondrial genome. The Amersham Kit (the original manufacturer's protocol) is not suitable for a DNA extraction from preserved specimens (and even not from control fresh/frozen tissue): PCR products were absent (Fig. 53(a)-Lanes: 30-40). The PCR results with modified protocol (A) were less successful than

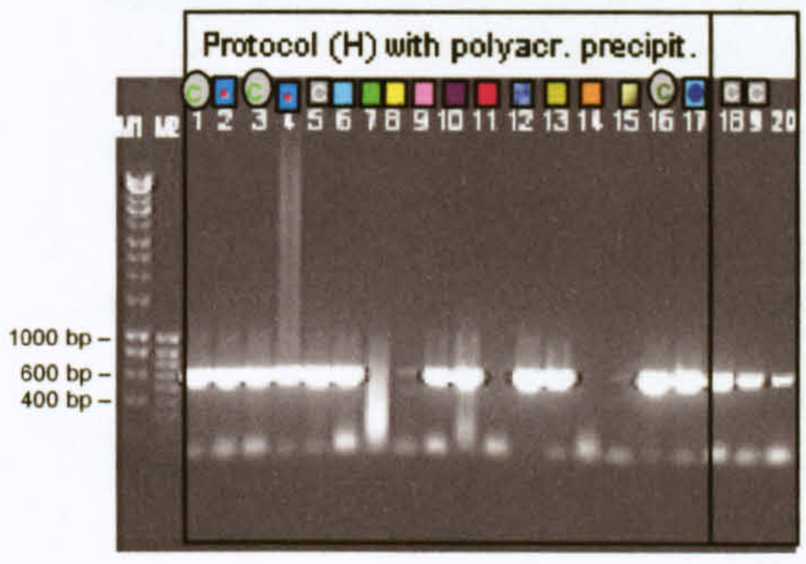
expected (Fig. 53(b)-Lanes: 31-43). This was probably caused by the modifications applied: liquid nitrogen was used instead of dry ice. Liquid nitrogen was unsuitable because it caused the powdered sample to be thrown out of the mortar and significantly reduced the amount of prepared tissue sample to be included in the DNA extraction procedure. Also, the washing/drying regime was much shorter (tissue samples were washed in 1xGTE buffer for only 24 hours and then 10-15 min washing in $sdH_2O$). The intrinsic factors related to a particular fish species might be also relevant (for example, mackerel is a much more "fatty" fish than *Nezumia* perhaps affecting the preservation of particular fishes, and the success in extracting PCR amplifiable arDNA with a particular DNA extraction protocol). The Qiagen and Promega Kits produced PCR-amplifiable arDNA that was extracted from pure ethanol and IMS preserved specimens, i.e. ethanol based preservative solutions. DNA from control fresh/frozen and few-days-ethanol preserved tissue samples was also amplifiable with the tested cyt b mitochondrial primer set.

PCR investigations with another mitochondrial gene (the 16S – Fig. 54) produced slightly different results than those with the cytochrome b gene (Fig. 53) using the same DNA extracts of mackerel. This supports the statements and previously shown evidence on the suitability of a particular DNA extraction protocol for amplifying a particular gene (i.e. genome region).

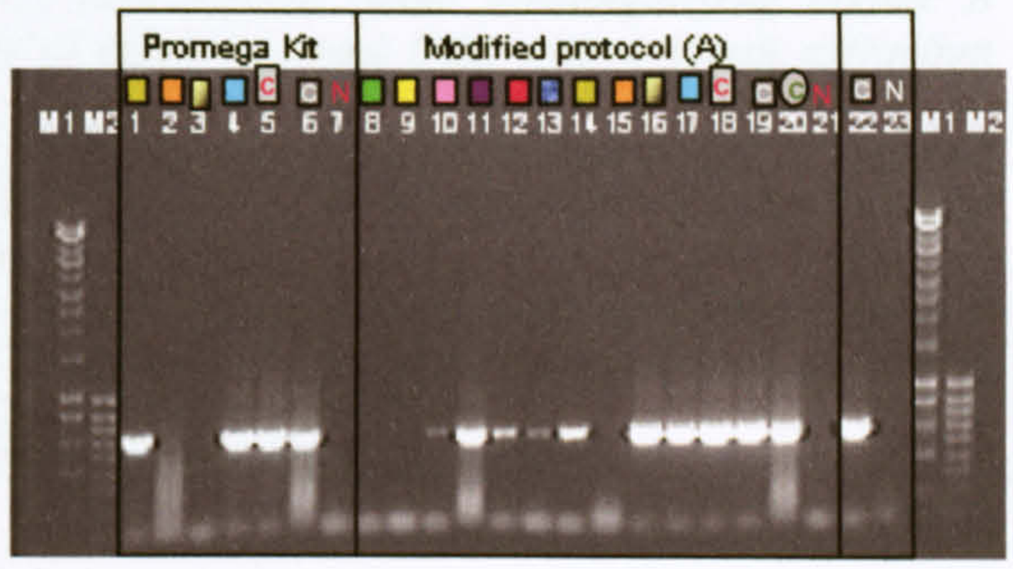
As in PCR experiments with cytochrome b (Fig. 53(a)-L: 1-12), the 16S mt-PCRs were also unsuccessful with any DNA extracts produced by protocol (H) that used 3M NaAc for DNA precipitation (Fig. 54(a)-Lanes: 1-26). Likewise, significant PCR success was observed if arDNA extracts were produced by protocol (H) with polyacrylamide instead of 3M NaAc for DNA precipitation (Fig. 54(a)-Lanes: 21-40; (b)-Lanes: 1-17). This modified protocol (H) was even successful for producing 3% unbuffered formalin-arDNA extract that generated a 16S product (faint band, but undoubtedly present) – Fig. 54(b)-Lane: 8). However, this was not the case in PCR experiments with cytochrome b. Only this DNA extraction protocol (from six tested) made it possible to obtain mt-PCR product with DNA extracted from 3% unbuffered formalin-preserved museum specimen of mackerel.



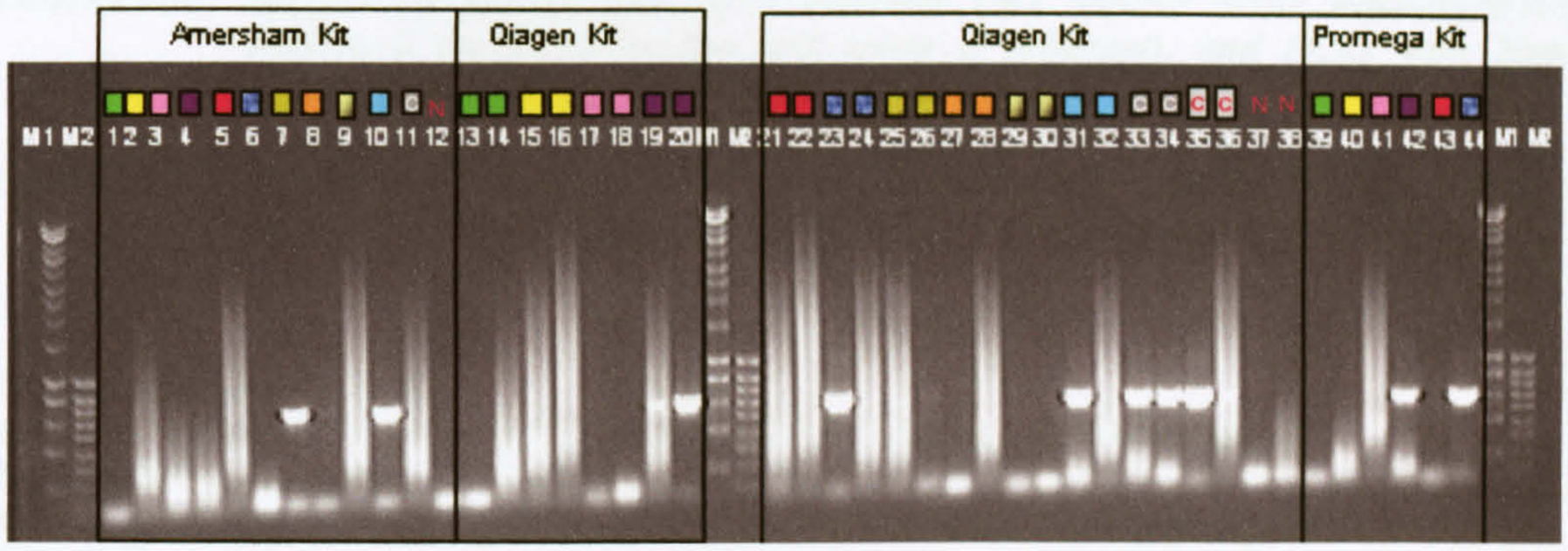

**(a)**




**(b)**




**(c)**




**(d)**

**Figure 54** – The 16S mitochondrial PCR amplifications with DNA of mackerel (*Scomber scombrus*) extracted by different DNA extraction protocols, using fresh/frozen and differently preserved muscle tissue samples. Muscle tissue was subjected to different washing/drying regimes prior to DNA extraction:

**A** - tissue was washed overnight in 1xTE buffer at room temperature prior to DNA extraction;
**B** - tissue was washed for a few minutes in $sdH_2O$, and then dried for a few minutes at room temperature prior to DNA extraction;
**C** - tissue was washed twice for 5 minutes in PBS (phosphate-buffered saline) and twice in $sdH_2O$ for a couple of minutes, and then briefly dried for a few minutes at room temperature prior DNA extraction;
**D** - tissue was washed in 1xGTE buffer for 24 hours (gently rotating the samples) at room temperature and frequently changing the buffer, then washed in $sdH_2O$ for 10-15 minutes, and then briefly dried for a few minutes at room temperature prior to DNA extraction.

**gel (a):**

**Lanes 1-26:** ***Protocol (H) - 3M NaAcetate (NaAc) DNA precipitation with applied tissue washing/drying regime A (indicated in the text below if different) – tail and dorsal muscle, and PCRs with blank extractions (Lanes: 25 and 26);***

**Lanes 27-40:** ***Modified protocol (H) - polyacrylamide DNA precipitation with applied tissue washing with 1xTE buffer at room temperature for various time (after 5, 4, 3, 2, 1.5, 1 and 0.5 hours of tissue washing with 1xTE buffer);***

**gel (b):**

**Lanes 1-17:** ***Modified protocol (H) - polyacrylamide DNA precipitation with applied tissue washing/drying regime A (indicated in the text below if different);***

**Lanes 18-20:** ***Promega Kit – Protocol (F) with applied tissue washing/drying regime B (indicated in the text below if different) – control fresh/frozen mackerel tissue, used 10x-100x diluted DNA extracts (10x-D to 100x-D) for PCRs;***

**gel (c):**

**Lanes 1-7:** ***Promega Kit – Protocol (F) with applied tissue washing/drying regime B (indicated in the text below if different), and PCR with a blank extraction (Lane: 7);***

**Lanes 8-21:** ***Modified protocol (A) – it was used liquid nitrogen instead of dry ice, applied tissue washing/drying regime D (indicated in the text below if different), and PCR with a blank extraction (Lane: 21);***

**Lane 22:** ***Promega Kit – Protocol (F) with applied tissue washing/drying regime B (indicated in the text below if different) – control fresh/frozen mackerel tissue, used 10x diluted DNA extract (10x-D) for a PCR;***

**Lane 23:** **PCR negative (sd$H_2O$ was used as PCR template) – "N"**

**gel (d):**

**Lanes 1-12:** ***Amersham Bioscience Kit with applied tissue washing/drying regime B (indicated in the text below if different), and PCR with a blank extraction (Lane: 12);***

**Lanes 13-38:** ***Qiagen Kit (tested both DNA elutions) with applied tissue washing/drying regime C (indicated in the text below if different), and PCRs with blank extractions (Lanes: 37 (I elution), 38 (II elution));***

**Lanes 39-44:** ***Promega Kit – Protocol (F) with applied tissue washing/drying regime B (indicated in the text below if different);***

- **10% unbuffered formalin**
  gel (a)-L: 1, 2; gel (b)-L: 7; gel (c)-L: 8; gel (d)-L: 1, 13 (I elution), 14 (II elution), 39
- **3% unbuffered formalin**
  gel (a)-L: 3, 4; gel (b)-L: 8; gel (c)-L: 9; gel (d)-L: 2, 15 (I elution), 16 (II elution), 40
- **50% isopropanol**
  gel (a)-L: 5, 6; gel (b)-L: 9; gel (c)-L: 10; gel (d)-L: 3, 17 (I elution), 18 (II elution), 41
- **70% IMS – industrial methylated spirits**
  gel (a)-L: 7, 8; gel (b)-L: 10; gel (c)-L: 11; gel (d)-L: 4, 19 (I elution), 20 (II elution), 42
- **Modified Steedman`s solution**
  gel (a)-L: 9, 10; gel (b)-L: 11; gel (c)-L: 12; gel (d)-L: 5, 21 (I elution), 22 (II elution), 43
- **70% ethanol**
  gel (a)-L: 11, 12; gel (b)-L: 12; gel (c)-L: 13; gel (d)-L: 6, 23 (I elution), 24 (II elution), 44
- **100% ethanol**
  gel (a)-L: 13, 14; gel (b)-L: 13; gel (c)-L: 1, 14; gel (d)-L: 7, 25 (I elution), 26 (II elution)
- **3% buffered formalin – with 4% sodium tetraborate**
  gel (a)-L: 15, 16; gel (b)-L: 14; gel (c)-L: 2, 15; gel (d)-L: 8, 27 (I elution), 28 (II elution)
- **10% buffered formalin – with 4% sodium tetraborate**
  gel (a)-L: 17, 18; gel (b)-L: 15; gel (c)-L: 3, 16; gel (d)-L: 9, 29 (I elution), 30 (II elution)
- **95% IMS**
  gel (a)-L: 19 (tail muscle), 20 (dorsal muscle); gel (b)-L: 6; gel (c)-L: 4, 17; gel (d)-L: 10, 31 (I elution), 32 (II elution)

- **95% IMS (tissue was washed for 0.5-5 hours in 1xTE buffer at room temperature prior to DNA extraction) – *Modified protocol (H) - polyacrylamide DNA precipitation*;**
gel (a)-L: 28 (5 h), 30 (4 h), 32 (3 h), 34 (2 h), 36 (1.5 h), 38 (1 h), 40 (30 min),
- **95% IMS (tissue was washed for 1-10 minutes in 1xTE buffer at room temperature prior to DNA extraction) – *Modified protocol (H) - polyacrylamide DNA precipitation*;**
gel (b)-L: 2 (10 min), 4 (1 min),
- **95% IMS (tissue was only dried at room temperature without any washing in 1xTE buffer or in other solution) – *Modified protocol (H) - polyacrylamide DNA precipitation*;**
gel (b)-L: 17,
- **fresh/frozen**
gel (a)-L: 21 (tail muscle), 23 (dorsal muscle); gel (b)-L: 5, 18 (10x-D), 19 (50x-D), 20 (100x-D); gel (c)-L: 6, 19, 22 (Promega Kit, 10-D); gel (d)-L: 11, 33 (I elution), 34 (II elution)
- **fresh/frozen (tissue was washed for 0.5-5 hours in 1xTE buffer at room temperature prior to DNA extraction) – *Modified protocol (H) - polyacrylamide DNA precipitation*;**
gel (a)-L: 27 (5 h), 29 (4 h), 31 (3 h), 33 (2 h), 35 (1.5 h), 37 (1 h), 39 (30 min),
- **fresh/frozen (tissue fresh/frozen (tissue was washed for 1-10 minutes in 1xTE buffer at room temperature prior to DNA extraction) – *Modified protocol (H) - polyacrylamide DNA precipitation*;**
gel (b)-L: 1 (10 min), 3 (1 min),
- **fresh/frozen (tissue was not washed at all, i.e. the tissue was used for DNA extraction straight from the freezer);**
gel (a)-L: 22 (tail muscle), 24 (dorsal muscle); gel (b)-L: 16; gel (c)-L: 20;
- **control 100% ethanol preserved – preserved in ethanol only for a few days**
gel (c)-L: 5, 18; gel (d)-L: 35 (I elution), 36 (II elution)

"N" – PCRs with blank extractions (extraction procedure performed without tissue, i.e. with sd$H_2O$ instead)
"M1" indicates quantitative size marker fragments – "Bioline" HyperLadder I
"M2" indicates quantitative size marker fragments – "Bioline" HyperLadder IV

The Amersham Kit arDNA extracts gave slightly better PCR performance with 16S than with cyt b, i.e. the 570 bp-region of the 16S mitochondrial gene was amplified with absolute ethanol-arDNA (Fig. 54(d)-Lane: 7) and with 95% IMS (industrial methylated spirits) arDNA of mackerel (Fig. 54(d)-Lane: 10). In previously described PCR experiments with cyt b (Fig. 53(a)-L: 30-40), PCR products were not detected with any of DNA extracts produced by the Amersham Kit (from preserved and from fresh/frozen tissue).

DNA extracts with the Promega Kit (protocol (F)) exhibited the same performance with both tested mitochondrial genes – PCR products were generated with DNA extracts from 70% IMS, 95% IMS, 70% and 100% pure ethanol preserved specimens, as well as with control DNA extracts from fresh/frozen and 100% pure ethanol preserved (tissue preserved in ethanol for only a few days) - Fig. 53(b) for cyt b; Fig. 54(c) and (d) for 16S.

The PCR performance of DNA extracts produced with the Qiagen Kit was similar to those produced with the Promega Kit (only IMS and pure ethanol arDNA extracts, and control DNAs generated 16S PCR products). The same PCR performance was exhibited with both tested mitochondrial genes: Fig. 53(a)-L: 41-45 and (b)-L: 1-18 for cyt b, and Fig. 54(d)-L: 13-38 for 16S).

The performance of DNA extracts produced by modified protocol (A) gave better results with the 16S gene (Fig. 54(c)-L: 8-21) than with the cytochrome b gene (Fig. 53(b)-L: 31-43): successful 16S-PCR amplification was obtained with 50% isopropanol-arDNA (Fig. 54(c)-L: 10), Steedman-arDNA (Fig. 54(c)-L: 12) and 70% pure ethanol-arDNA (Fig. 54(c)-L: 13). This was not the case with the cyt b gene (Fig. 53(b)). This was the only protocol which produced arDNA extract from Steedman-preserved tissue that generated PCR product. This is supporting evidence that protocol (A) developed in the main study with *Nezumia* was the appropriate protocol for this type of preserved sample. Also, this was the only protocol that produced arDNA extract from 10% buffered tissue samples of mackerel that gave successful PCR products with both tested mitochondrial genes (Fig. 53(b)-Lane: 39 for cyt b, and Fig. 54(c)-Lane: 16 for 16S). As already mentioned, the overall performance of protocol (A) would be probably have been better in experiments with mackerel if the described modifications of protocol (A) had not been applied (shorter period of washing a tissue sample prior to DNA extraction and the use of liquid nitrogen instead of dry ice).

It was surprising that this modified protocol (A) produced 70% and 100% pure ethanol arDNA extracts of a low mitochondrial PCR suitability – products were absent (with 70% ethanol-arDNA, cyt b; Fig. 53(b)-Lane: 36) or of a poor yield (Fig. 53(b)-Lane: 37 with cyt b; Fig. 54(b)-Lane: 13). However, arDNA extracts from 70% and 95% IMS preserved specimens gave better PCR performance, especially with the 16S mitochondrial gene. This might suggest that this protocol is not the most suitable DNA extraction protocol for producing PCR amplifiable mitochondrial arDNA from pure ethanol preserved fish specimens, but it is too early to draw this kind of conclusion because arDNA extracts from ethanol preserved tissue of *Coryphaenoides* produced very good mt-PCR results (e.g. Fig. 49(b)-Lanes: 1 and 2 for 16S; Fig. 51(a)-Lanes: 13-17, (c)-Lanes: 1-5 for COIII).

None of the tested DNA extraction protocols produced arDNAs from the 3% buffered formalin preserved specimen of mackerel that were suitable for cyt b and/or 16S mitochondrial amplifications. The explanation might lie in the poorly preserved tissue of this fish specimen (morphologically, the tissue was not firm, and the fish was soft and floppy). The preservation fluid of this specimen (3% buffered formalin) was opaque and flocculent. This might suggest tissue degeneration with a likely negative effect on DNA integrity. Also, this might suggest that the use of a lower concentration of chemical/s (in this case, formalin) is not always the best solution regarding recovery of DNA from preserved

samples. In support of this are the successful mt-PCR amplifications with 10% buffered formalin-arDNA if arDNA extracts were produced by protocol (A). Unbuffered formalin exhibited an extremely negative effect on recovery of DNA, especially if used as a 10% preservation solution (only one arDNA extract [produced by protocol (H) – polyacrylamide precipitation] generated a weak 16S PCR product). These preliminary PCR experiments on arDNA of 4-yrs preserved mackerel specimens suggest that the choice and concentration of chemicals used in preservation is of a crucial importance for molecular work and recovery of DNA, much more significant than the application of a particular DNA extraction protocol.

How much intrinsic factors of the fish species (or species of any other group of organisms) are responsible for variations in producing PCR amplifiable arDNA (in general and for a specific marker system) is a question that still remains unanswered. Further investigations on a range of fresh/frozen and differently preserved fish (and other) species should provide a clearer answer to this question, as well as answers about the specificity, accessibility and recovery of mitochondrial arDNA in preserved specimens of different species that have different size, structure and organisation of mitochondrial genomes.

If fresh/frozen tissue samples of mackerel were exposed for a prolonged time to room temperature, in a buffer and/or $sdH_2O$ for tissue washing, it might have a negative effect on the recovery of DNA from fresh/frozen tissue. This depended on buffer used and length of exposure of the tissue sample to room temperature. For successful mt-PCR amplifications with cytochrome b, it was noticed that the buffer used for eliminating the excess of preservative was the major factor. If 1xTE buffer was used, the length of washing fresh/frozen tissue at room temperature did not affect successful recovery of mtDNA from tissue (Fig. 53(a)-Lanes: 13, 15, 17, 28). However, if 1xGTE buffer and $sdH_2O$ were used, the length of exposure of fresh/frozen tissue to room temperature (and to this solution) had a significant adverse effect on recovery of mtDNA (Fig. 53(b)-Lanes: 42, 43). This experiment indicates the importance of handling caught animals intended to be used for archival collections and molecular work.

### 3.3.2. Comparison of relevant findings by other researchers and results in this study: discussion on mtDNA from preserved specimens

The exact reason(s) for the difficulties in amplifying mitochondrial sequences from formalin- and Steedman's-preserved fishes is not clear, but other researchers have also reported difficulties in amplifying formalin-mtDNA from fishes and other preserved organisms. Some studies reported that it was possible to amplify only short PCR fragments of mtDNA (usually 100-300 bp) using formalin-fixed specimens (for example, Bucklin and Allen (2004) on zooplankton stored for a long period in buffered formalin; Boyle *et al.* (2004) and Zardus *et al.* (2006) on a formalin-fixed, ethanol preserved deep-sea bivalve). Rohland *et al.* (2004) experienced difficulties in amplifying mitochondrial genes using DNA extracted from dry museum specimens of chimpanzees and hyenas, even after increasing the number of the PCR cycles up to 60. Collins *et al.* (2002) were unable to successfully amplify and sequence the 16S mitochondrial region of swamp eels (*Monopterus albus* and *M. cuchia*) from formalin-fixed, ethanol-preserved specimens stored for 5, 20 or 30 years in the archival collection of the University of Michigan Museum of Zoology. They did not have a problem in amplifying and sequencing the 615-bp portion of the 16S mitochondrial gene if fresh/frozen and 70-95% ethanol-preserved fish specimens were used, but they did not succeed with the formalin-fixed samples even when using internally designed primers that should amplify only a 250-375 base-pair region of the 16S gene. The study of Whittier *et al.* (1999) also reported that a decrease of PCR product sizes from ~418 to ~245 bp for the COII gene, or to 100-200 bp for COIII gene (Diaz-Viloria *et al.* 2005), did not improve mitochondrial amplification results with formalin-arDNA. It seems that the size of targeted mitochondrial DNA regions is unlikely to be always of crucial importance in amplifying mitochondrial gene sequences from preserved specimens (see Junqueira *et al.* 2002), although smaller fragments are easier to amplify according to some researchers (Wirgin *et al.* 1997; Boyle *et al.* 2004). Diaz-Viloria *et al.* (2005) and Chakraborty *et al.* (2006) generally question the integrity of mtDNA in formalin-fish specimens that have been stored for long periods.

According to the findings of Banerjee and Brown (2002) on archaeological plant remains, chloroplast DNA exhibited faster degradation than nuclear DNA. They successfully amplified nuclear glutenin genes, but repeated attempts to amplify the chloroplast rbcL locus were unsuccessful. Also, they found that storage of leaf material for up to 289 days at ambient temperature resulted in no significant degradation of nuclear DNA but extensive

degradation of chloroplast DNA. Banerjee and Brown (2002) concluded that a differential preservation of nuclear and chloroplast DNA is probably caused by differences in the biochemical environments in the nuclei and in the chloroplast. They suggested that reactive oxygen such as superoxide radicals ($O_2^-$) and hydrogen peroxide ($H_2O_2$), which are derived from molecular oxygen ($O_2$), are potent agents of DNA damage in the plastid.

The similarity between mitochondria and chloroplasts from the aspect of energy production and oxidative processes is evident. It is known that mitochondria generate free oxygen radicals causing an internal environment with a high mutagenic potential (Balaban *et al.* 2005), as well as proven faster degradation of mtDNA vs. nDNA in frozen samples (Berger *et al.* 2001). Could the findings of Banerjee and Brown (2002) on chloroplast DNA damage caused by reactive oxygen and differential preservation of nuclear and organelle (chloroplast) DNA also explain a faster degradation of mitochondrial DNA and much lower presence of mtDNA in formalin-fixed/preserved organisms, or, is it just more difficult to access mtDNA from preserved organisms? Is this related to the size and specific organisation (architecture) of mitochondrial genomes in different species and groups of organisms (invertebrates and vertebrates in particular)? These questions could be answered only by adequate investigations on mitochondria and mitochondrial DNA of a range of preserved and non-preserved organisms, but the results from this study indicate limitations and difficulties in using mitochondrial genes for molecular investigations in formalin/Steedman's fish specimens (mackerel specimens stored for 4 years and *Nezumia* specimens stored for up to 20 years).

Whittier *et al.* (1999) are of the opinion that formalin-mtDNA versus formalin-nDNA might be more susceptible for hydroxymethylation and inhibition during PCR amplification because of more abundant AT-rich regions in mitochondrial DNA than in nuclear DNA. If the suggestions of Chang and Loew (1994) are correct that formaldehyde causes hydroxymethylation of exposed exocyclic amino groups and that the A-T rich regions are more susceptible to oxidative attack compared to G-C base pairs, the explanation of Whittier *et al.* (1999) might be acceptable as one of the reasons for increased difficulties in PCR amplifications of mtDNA vs. nDNA (see also Tang (2006) and Skage and Schander (2007)). However, the cause is probably much more complex than this single explanation related exclusively to mtDNA as being adenine-thymine base pairs (A-T) rich genome (especially given that in this study the problem about amplifying and sequencing 458 bp A-T rich region from an unknown part of genome was not an issue; see section 3.2).

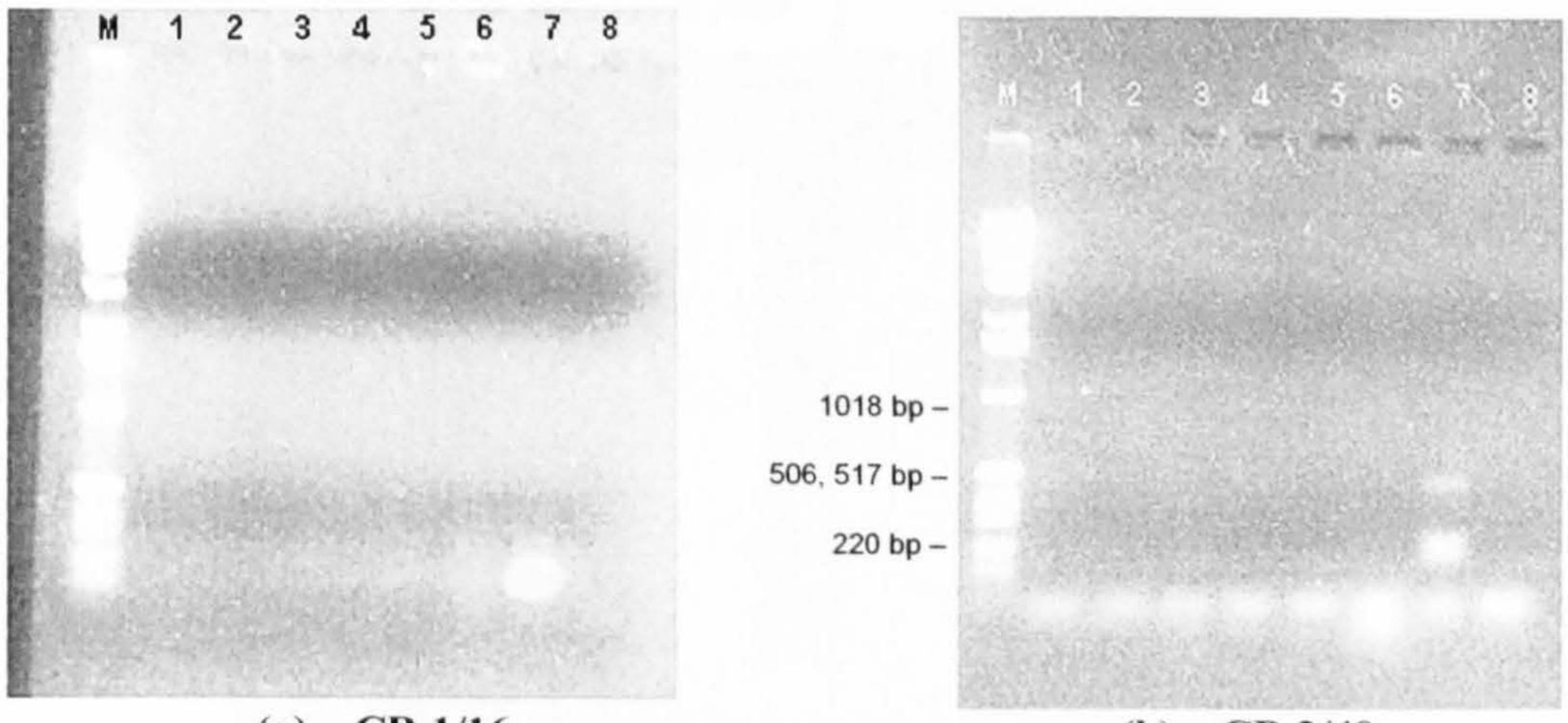
Different structure, base composition and organisation of the mitochondrial genome (and/or specific regions) in different species and groups of organisms (invertebrates and vertebrates in particular) might be important for the integrity of mtDNA and/or accessibility of mtDNA (see Boore *et al.* 2005; Chakraborty *et al.* 2006), but this is not clear due to insufficient research on this aspect. Gilbert *et al.* (2005b) suggest that structural elements of mitochondrial DNA confer a degree of *in vivo* and post-mortem protection from sequence modification. However the problems related to extracting PCR amplifiable mt-arDNA might also be related to the way of handling the specimens when they are caught (for how long they were dead without being fixed?) and much faster degradation of mtDNA than nDNA (Berger *et al.* 2001). These aspects should be taken into consideration if records are available for given collections. The reaction of formaldehyde with DNA is known to be damaging, especially over long periods of time, but understanding of DNA-formalin interactions, kinds of DNA damage and their distribution in genomes still remains largely incomplete (Tang 2006; Wandeler *et al.* 2007). More and more researchers are of the opinion that mtDNA from formalin-fixed samples is less accessible than is nuclear DNA (see Tang 2006).

The type of tissue (brain, muscle, heart etc.) and kind of sample (faeces, hair, urine etc.) used in DNA extraction might also influence the success of mitochondrial PCR amplifications (Whittier *et al.* 1999; Miething *et al.* 2006). The results of this study did not show a significant correlation between type of tissue and the success of mt-PCR amplifications, although more extensive investigations are required before drawing definite conclusions. The success of mt-PCR amplification in this study was mostly related to a particular arDNA extract and DNA extraction protocol. Supplementary experiments of this study, carried out on ten differently preserved mackerel specimens, suggest that preservation of fish specimens is the most important factor in ability to amplify mitochondrial DNA sequences, but the DNA extraction protocol used is also significant. The crucial question is: "Is it a problem of faster degradation of mtDNA, or of inaccessibility of mtDNA from preserved specimens?" The results from this study cannot clearly provide the answer, but it seems that it is a combination of both. Difficulty in amplifying mitochondrial genome regions from formalin-arDNA of some species and groups of organisms might prevent the application of mitochondrial barcodes (16S and COI) to all preserved specimens. There are already suggestions for using nuclear ribosomal genes for barcoding purposes (Floyd *et al.* 2002; Bhadury *et al.* 2006b; Chu *et al.* 2006), but applicability to all, or the majority of, preserved specimens and species remains to be assessed.

## 3.4. Other PCR experiments

A limited number of PCR experiments with **ITS primers** did not generate any PCR products with **formalin-arDNA of *Nezumia***. These should not be regarded as definitive results on *Nezumia* species, because the PCRs were performed only with a few formalin-arDNA extracts of *Nezumia* and without PCR optimisation for this molecular marker.

The PCR experiments with **microsatellites** were carried out with the expectation that primers designed for *Coryphaenoides* (by Dr Alex Rogers's research team in Southampton) might be applicable to the closely related species of *Nezumia*. The *Coryphaenoides* microsatellite primer sets CR 1/16 and CR 2/40 applied on **formalin-arDNA of *Nezumia aequalis*** and ***N. micronychodon*** did not produce PCR products (Fig. 55(a) and (b)-Lanes: 1-5). Also, PCR products were not generated with DNA of cod (Fig. 55(a) and (b)-L: 6). However, PCR amplifications with ethanol-arDNA of *Coryphaenoides rupestris* generated strong and clear bands (PCR products) of the appropriate size (100 bp – 150 bp) - Fig. 55(a) and (b)-L: 7. It might be worthwhile to test these microsatellite primers again against DNA of *Nezumia*, but with DNA of a good quality (from fresh/frozen and/or short-term ethanol preserved samples).




**(a) – CR 1/16** **(b) – CR 2/40**

**Figure 55** – PCR amplifications with the microsatellite primer set CR 1/16 **(a)** and set CR 2/40 **(b)**. *Nezumia aequalis* (L: 1, 2) and *Nezumia micronychodon* formalin-arDNA: Lanes 3-5 on gels (a) and (b). Cod DNA – L: 6 on gels (a) and (b). Ethanol-arDNA of *Coryphaenoides rupestris*: Lane 7 on (a) and (b).
"M" indicates size marker fragments - 1 Kb DNA Ladder.

**Additional experiments** on control, ethanol preserved *Nezumia* specimens and supplementary experiments on differently preserved fish specimens of mackerel were performed in order to investigate the possibility of amplifying a single/low-copy and multi-copy targeted sequences of preserved specimens (regions that belong to different nuclear genes); that is, the feasibility of amplifying **rhodopsin** and **28S ribosomal genes**. The experiments on mackerel were carried out in order to investigate the effect of different preservation methods and the application of different DNA extraction protocols on the ability to use different molecular marker systems for extracting molecular information from preserved fish specimens.

### 3.4.1. Control, ethanol preserved *Nezumia*

The preliminary PCR investigations on short-term (1.5-2 yrs) ethanol preserved specimens of *Nezumia* with the **28S gene** gave promising results about the possibility to use this primer set on *Nezumia* species. Generated PCR products have sizes of about 1000 bp (Fig. 56). This primer set developed by Lecointre *et al.* (1997) for fish is definitely applicable to *N. aequalis* and *N. micronychodon* species for amplifying this 28S gene D2 domain and can potentially be used for molecular investigation on *Nezumia*.

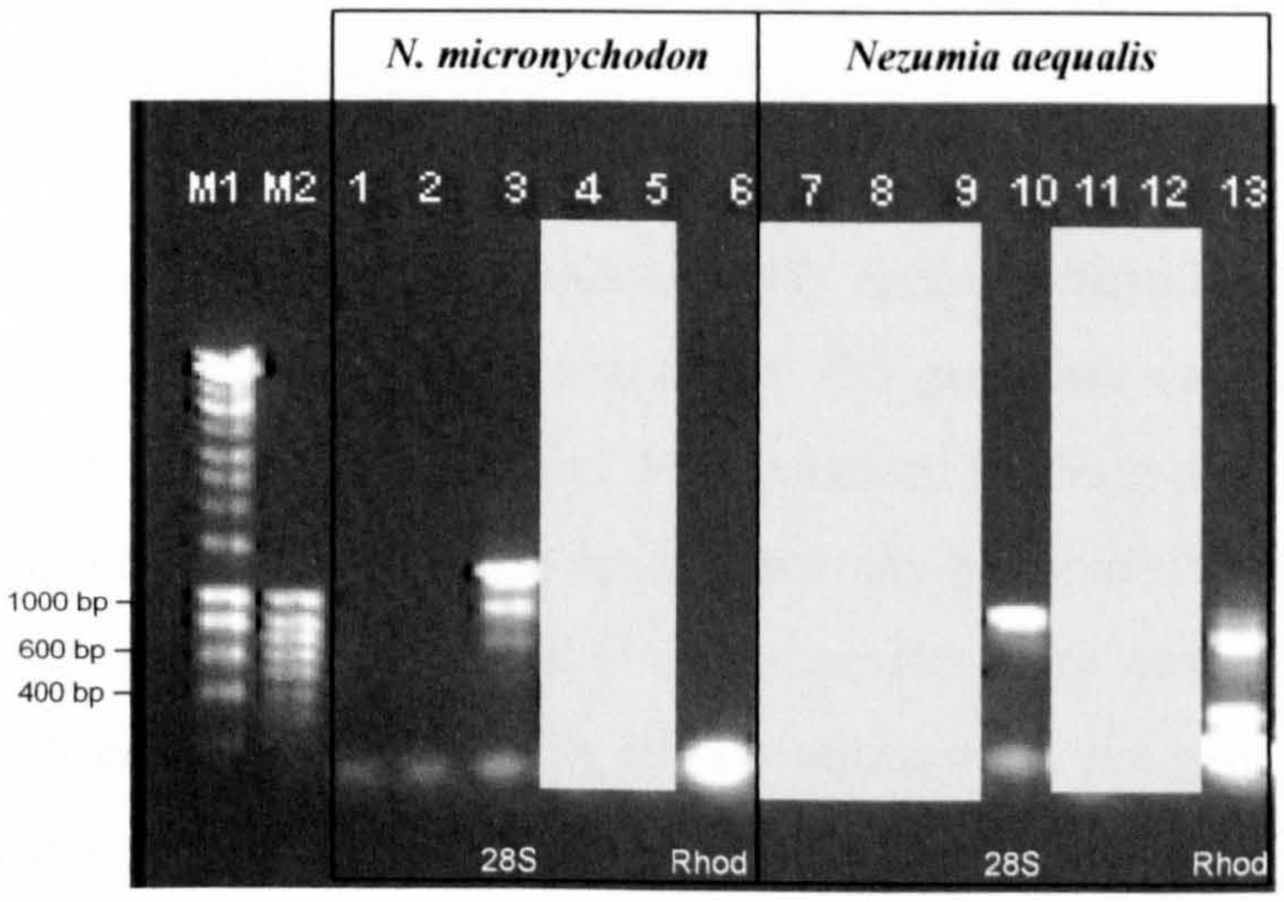



**Figure 56** – PCR amplifications of the ***28S gene*** (L: 3 and 10) and the ***rhodopsin gene*** (L: 6 and 13) with ethanol-arDNA of *Nezumia micronychodon* and *Nezumia aequalis* (tissue specimens were preserved only for 1.5 - 2 years).
"M1" indicates quantitative size marker fragments – "Bioline" HyperLadder I
"M2" indicates quantitative size marker fragments – "Bioline" HyperLadder IV

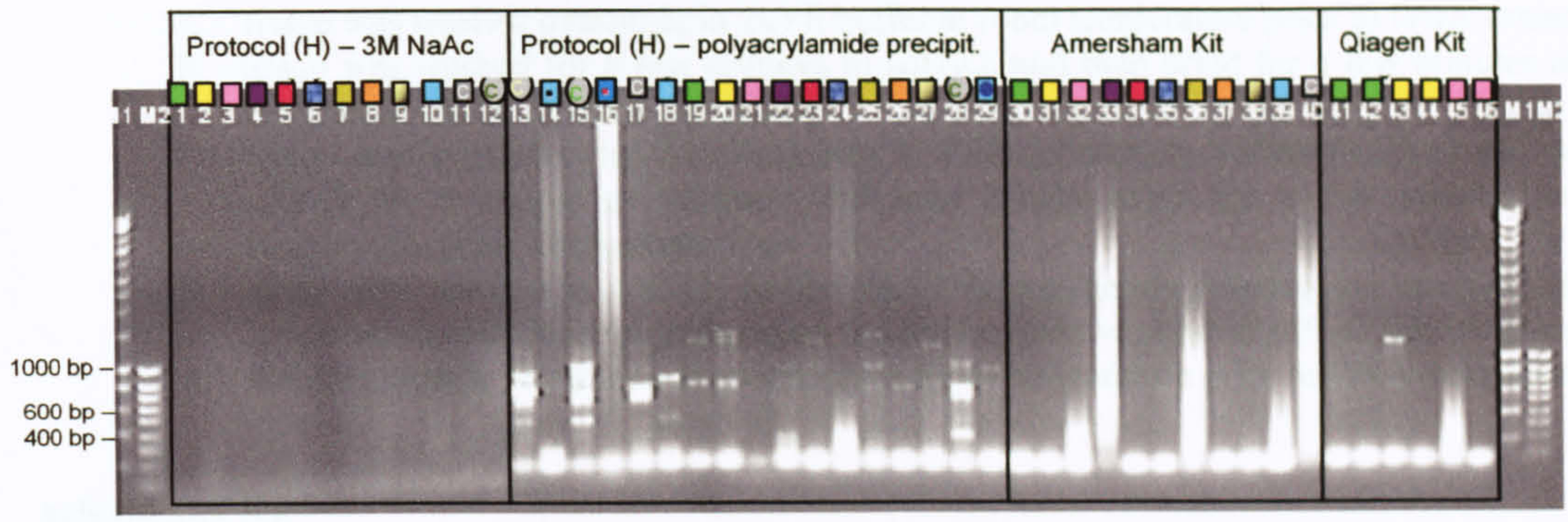
A region of the **rhodopsin** gene (single/low copy gene) was amplified with ethanol-arDNA of *Nezumia aequalis* (Fig. 56, Lane 13), but not with ethanol-arDNA of *N. micronychodon* (Fig. 56, Lane 6). These are preliminary results and without confirmation of these results on other fish specimens (fresh/frozen in particular) it is too early to draw any conclusion about the usability of this primer set on one species but not on another.

### 3.4.2. Supplementary evidence on differently preserved specimens of mackerel

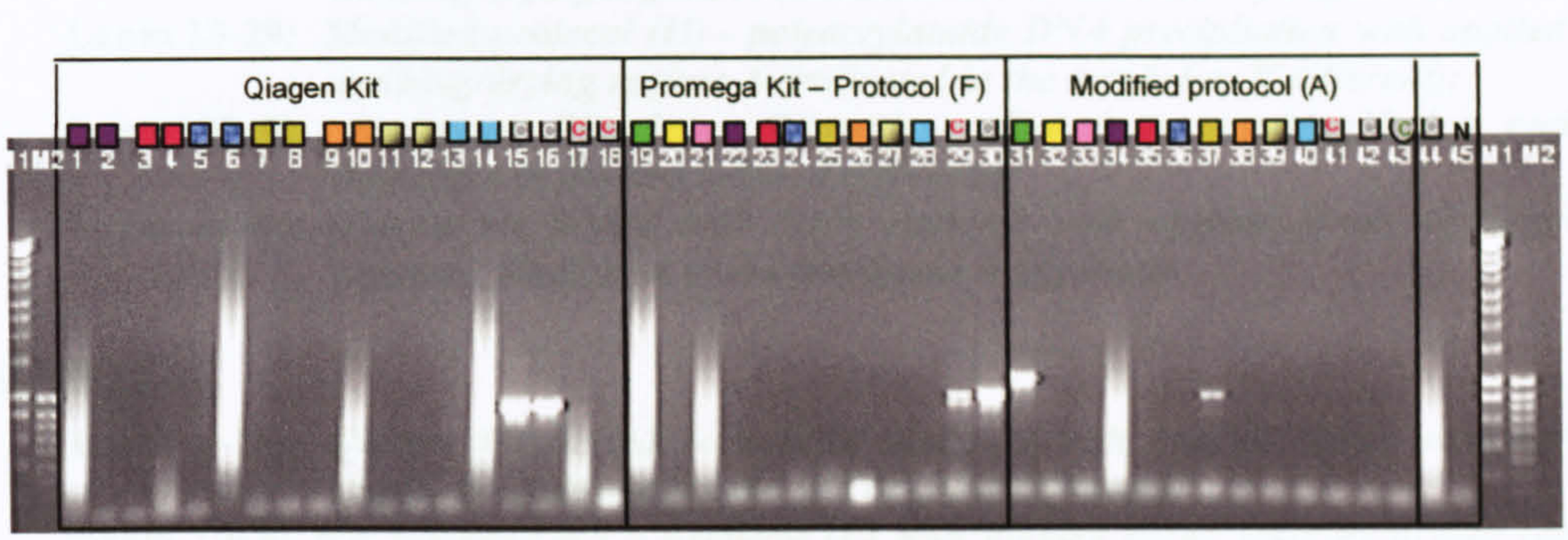
Overall, PCR amplifications with 28S (many-copy) and rhodopsin (a single/low-copy) genes on ten differently preserved specimens of mackerel (Fig. 57) were less successful than with mitochondrial genes (Figs. 53 and 54).

The amplification success with the **28S gene** depended on the type of preservation and DNA extraction protocol applied. The same situation occurred as with mitochondrial amplifications; the PCR products were lacking on all mackerel specimens (even fresh-frozen one) if the phenol protocol (H) with 3M NaAc precipitations was used in DNA extraction procedure (Fig. 57(a)-L: 1-12). Similarly, if the same DNA extraction protocol (H), but with polyacrylamide DNA precipitation, was applied, the PCR amplification results were more successful (Fig. 57(a)-L: 13-29).

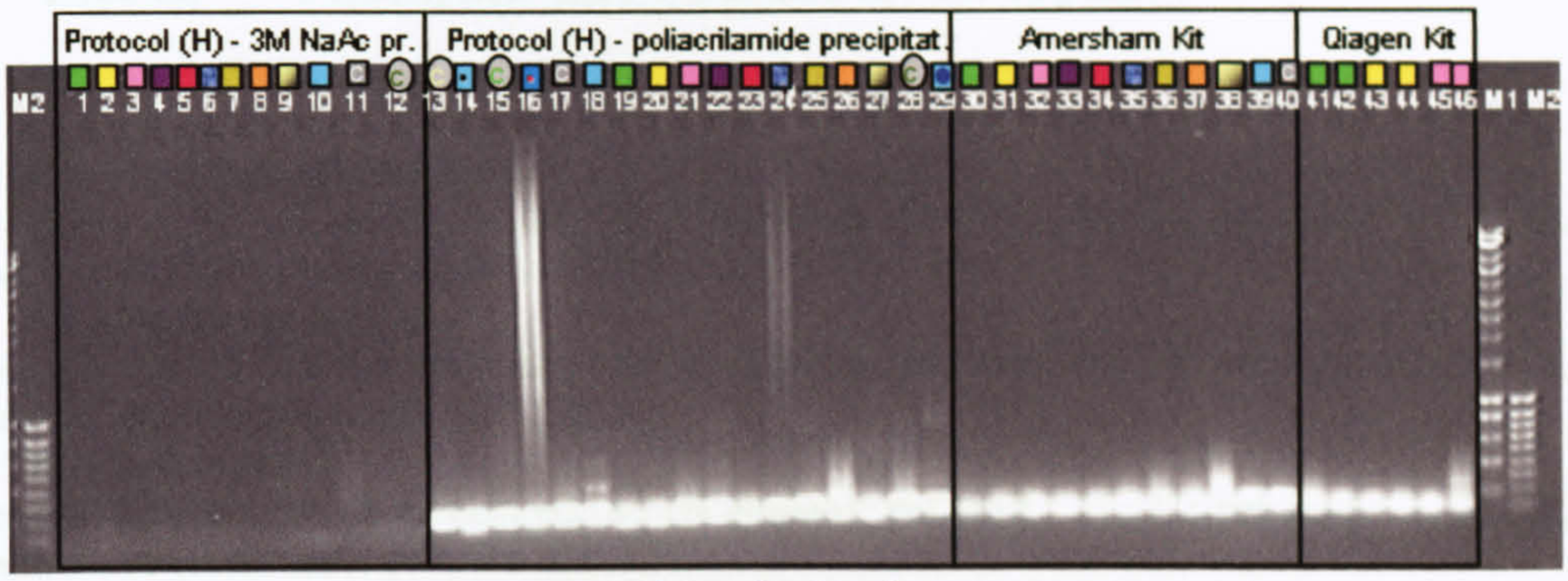
The arDNA extracts produced by protocol (H) (polyacrylamide DNA precipitation) exhibited some specificity in amplification of the 28S gene that was not characteristic for amplifications with mitochondrial genes. For example, washing/drying regimes of 95% IMS tissue samples seemed to be much more important in amplifying this 28S gene than it was for amplifying mitochondrial genes. If tissue samples were washed in 1xTE buffer for shorter periods (5 hours, or few minutes), PCR products were not generated (Fig. 57(a)-L: 14 and 16). However, if 95% IMS preserved tissue samples were washed overnight in 1xTE buffer, or just dried out without any tissue pre-washing, PCR products were generated with this 28S primer set (Fig. 57(a)-L: 18 and 29) – the bands were stronger if tissue was pre-washed in 1xTE buffer than just dried out.



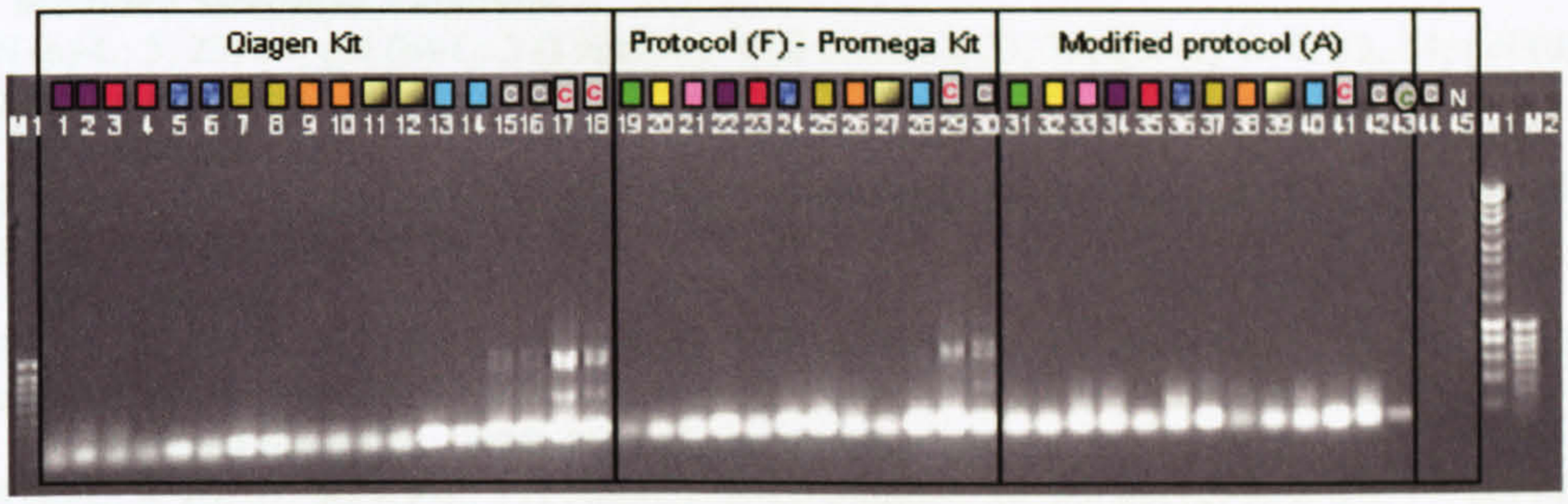

**(a) – 28S gene**




**(b) – 28S gene**




**(c) – rhodopsin gene**




**(d) – rhodopsin gene**

**Fig. 57** – The ***28S gene*** PCR amplifications **(a) and (b)**, and ***the rhodopsin gene*** **(c) and (d)**, PCR amplifications with DNA of mackerel (*Scomber scombrus*) extracted by different DNA extraction protocols, using fresh/frozen and differently preserved muscle tissue samples. Muscle tissue was subjected to different washing/drying regimes prior to DNA extraction:

**A** - tissue was washed overnight in 1xTE buffer at room temperature prior to DNA extraction;
**B** - tissue was washed for a few minutes in $sdH_2O$, and then dried for a few minutes at room temperature prior to DNA extraction;
**C** - tissue was washed twice for 5 minutes in PBS (phosphate-buffered saline) and twice in $sdH_2O$ for a couple of minutes, and then briefly dried for a few minutes at room temperature prior DNA extraction
**D** - tissue was washed in 1xGTE buffer for 24 hours (gently rotating the samples) at room temperature and frequently changing the buffer, then washed in $sdH_2O$ for 10-15 minutes, and then briefly dried for a few minutes at room temperature prior to DNA extraction

**gels (a) and (c):**

**Lanes 1-12:** ***Protocol (H) - 3M NaAcetate (NaAc) DNA precipitation with applied tissue washing/drying regime A (indicated in the text below if different);***
**Lanes 13-29:** ***Modified protocol (H) - polyacrylamide DNA precipitation with applied tissue washing/drying regime A (indicated in the text below if different);***
**Lanes 30-40:** ***Amersham Bioscience Kit with applied tissue washing/drying regime B (indicated in the text below if different);***
**Lanes 41-46:** ***Qiagen Kit (tested both DNA elutions) with applied tissue washing/drying regime C (indicated in the text below if different);***

**gels (b) and (d):**

**Lanes 1-18:** ***Qiagen Kit (tested both DNA elutions) with applied tissue washing/drying regime C (indicated in the text below if different);***
**Lanes 19-30, 45:** ***Promega Kit – Protocol (F) with applied tissue washing/drying regime B (indicated in the text below if different);***
**Lanes 31-43:** ***Modified protocol (A) – it was used liquid nitrogen instead of dry ice, applied tissue washing/drying regime D (indicated in the text below if different);***
**Lane 45:** ***PCR negative ($sdH_2O$ was used as PCR template)***

- **10% unbuffered formalin**
  gel (a)-L: 1, 19, 30, 41 (I elution), 42 (II elution); gel (b)-L: 19, 31; gel (c)-L: 1, 19, 30, 41 (I elution), 42 (II elution); gel (d)-L: 19, 31
- **3% unbuffered formalin**
  gel (a)-L: 2, 20, 31, 43 (I elution), 44 (II elution); gel (b)-L: 20, 32; gel (c)-L: 2, 20, 31, 43 (I elution), 44 (II elution); gel (d)-L: 20, 32
- **50% isopropanol**
  gel (a)-L: 3, 21, 32, 45 (I elution), 46 (II elution); gel (b)-L: 21, 33; gel (c)-L: 3, 21, 32, 45 (I elution), 46 (II elution); gel (d)-L: 21, 33
- **70% IMS – industrial methylated spirits**
  gel (a)-L: 4, 22, 33; gel (b)-L: 1 (I elution), 2 (II elution), 22, 34; gel (c)-L: 4, 22, 33; gel (d)-L: 1 (I elution), 2 (II elution), 22, 34
- **Modified Steedman`s solution**
  gel (a)-L: 5, 23, 34; gel (b)-L: 3 (I elution), 4 (II elution), 23, 35; gel (c)-L: 5, 23, 34; gel (d)-L: 3 (I elution), 4 (II elution), 23, 35
- **70% ethanol**
  gel (a)-L: 6, 24, 35; gel (b)-L: 5 (I elution), 6 (II elution), 24, 36; gel (c)-L: 6, 24, 35; gel (d)-L: 5 (I elution), 6 (II elution), 24, 36
- **100% ethanol**
  gel (a)-L: 7, 25, 36; gel (b)-L: 7 (I elution), 8 (II elution), 25, 37; gel (c)-L: 7, 25, 36; gel (d)-L: 7 (I elution), 8 (II elution), 25, 37
- **3% buffered formalin – with 4% sodium tetraborate**
  gel (a)-L: 8, 26, 37; gel (b)-L: 9 (I elution), 10 (II elution), 26, 38; gel (c)-L: 8, 26, 37; gel (d)-L: 9 (I elution), 10 (II elution), 26, 38
- **10% buffered formalin – with 4% sodium tetraborate**
  gel (a)-L: 9, 27, 38; gel (b)-L: 11 (I elution), 12 (II elution), 27, 39; gel (c)-L: 9, 27, 38; gel (d)-L: 11 (I elution), 12 (II elution), 27, 39
- **95% IMS**

gel (a)-L: 10 (3M NaAc), 18 (polyacryl.), 39; gel (b)-L: 13 (I elution), 14 (II elution), 28, 40; gel (c)-L: 10 (3M NaAc), 18 (polyacryl.), 39; gel (d)-L: 13 (I elution), 14 (II elution), 28, 40

- **95% IMS (tissue was was washed for 0.5-5 hours in 1xTE buffer at room temperature prior to DNA extraction) – *Modified protocol (H) - polyacrylamide DNA precipitation*;**
gel (a)-L: 14 (5 h); gel (c)-L: 14 (5 h)
- **95% IMS (tissue was washed for 1-10 minutes in 1xTE buffer at room temperature prior to DNA extraction) – *Modified protocol (H) - polyacrylamide DNA precipitation*;**
gel (a)-L: 16 (1 min); gel (c)-L: 16 (1 min),
- **95% IMS (tissue was only dried at room temperature without any washing in 1xTE buffer or in other solution) – *Modified protocol (H) - polyacrylamide DNA precipitation*;**
gel (a)-L: 29; gel (c)-L: 29
- **fresh/frozen**
gel (a)-L: 11, 17, 40; gel (b)-L: 15 (I elution), 16 (II elution), 30 (10x-D), 42, 44 (Promega Kit; 100x-D); gel (c)-L: 11, 17, 40; gel (d)-L: 15 (I elution), 16 (II elution), 30 (10x-D), 42, 44 (Promega Kit; 100x-D)
- **fresh/frozen (tissue tissue was washed for 0.5-5 hours in 1xTE buffer at room temperature prior to DNA extraction) – *Modified protocol (H) - polyacrylamide DNA precipitation*;**
gel (a)-L: 13 (5 h); gel (c)-L: 13 (5 h)
- **fresh/frozen (tissue was washed for 1-10 minutes in 1xTE buffer at room temperature prior to DNA extraction) – *Modified protocol (H) - polyacrylamide DNA precipitation*;**
gel (a)-L: 15 (1 min); gel (c)-L: 15 (1 min)
- **fresh/frozen (tissue was not washed at all, i.e. the tissue was used for DNA extraction straight from the freezer);**
gel (a)-L: 12, 28; gel (b)-L: 43; gel (c)-L: 12, 28; gel (d)-L: 43
- **control 100% ethanol preserved – preserved in ethanol only for a few days**
gel (b)-L: 17 (I elution), 18 (II elution), 29, 41; gel (d)-L: 17 (I elution), 18 (II elution), 29, 41

"M1" indicates quantitative size marker fragments – "Bioline" HyperLadder I
"M2" indicates quantitative size marker fragments – "Bioline" HyperLadder IV

Interestingly, PCR products (faint bands) were obtained with formalin-arDNA extracts – 3% and 10% unbuffered and buffered formalin (Fig. 57(a)-L: 19, 20, 26 and 27) if protocol (H) with polyacrylamide was applied. The 70% IMS and 70% ethanol-arDNA extracts did not generate 28S PCR products. This was contradictory to the amplification success of mitochondrial genes with these arDNA extracts (Fig. 53(a)-L: 22 and 24 for cyt b, and Fig. 60(b)-L: 10 and 12 for 16S). There were some faint bands in amplifying the 28S gene with 70% IMS-arDNA of mackerel (Fig. 57(a)-L: 22), but not with 70% pure-ethanol arDNA extract (Fig. 57(a)-L: 24). Some faint bands also were generated with Steedman's-arDNA (Fig. 57(a)-L: 23). PCR products (28S) were generated in all PCR amplifications if DNA extracts from fresh/frozen tissue samples were applied (Fig. 57(a)-L: 13, 15, 17 and 28).

Archival DNA extracts produced by the Amersham Kit did not generate any PCR products with the 28S gene (Fig. 57(a)-L: 30-40), not even with DNA extracts from control fresh/frozen samples (Lane 40). A similar situation was also recorded with PCR amplification of mitochondrial genes. It seems that this Kit is not useful for extracting DNA from preserved specimens. Because of the low performance of this Kit, even with

control (unpreserved) specimens, it is a questionable if it was something wrong with this particular kit-box, or if the Amersham Kit is not suitable for extracting DNA from fish specimens (preserved and unpreserved) in general.

Archival DNA extracts produced by Qiagen Kit (Fig. 57(a)-L: 41-46, (b)-L: 1-18), Promega Kit (Fig. 57(b)-L: 19-30), and modified protocol (A) - (Fig. 57(b)-L: 31-43) were less successful in amplifying the 28S nuclear region than in amplifying targeted mitochondrial sequences (Figs. 53 and 54) and RAPDs (data not shown). The 28S PCR products were generated with 3% unbuffered formalin-arDNA produced by Qiagen Kit – only with the first DNA elution (Fig. 57(a)-L: 43), and with DNA from fresh/frozen tissue – both DNA elutions (Fig. 57(b)-L: 15 and 16). DNA extracts produced by the Promega Kit gave successful 28S PCR amplifications only with control samples, i.e. DNA extracted from fresh/frozen and stored for a few days in absolute ethanol-preserved tissue samples (Fig. 57(b)-L: 29 and 30). For DNA extracts produced by protocol (A), only 10% unbuffered formalin-arDNA extract (Fig. 57(b)-L: 31) and absolute, pure-ethanol-arDNA extract (Fig. 57(b)-L: 37) generated 28S PCR products. The 28S PCR product generated from 10% unbuffered formalin-arDNA (Fig. 57(b)-L: 31) gave a larger sized PCR product than other "successful extracts". Sequencing of PCR products should reveal the true nature of this bigger product, i.e. if it was a PCR artifact.

A low amplification success with the ~1000 bp region of 28S gene with mackerel arDNA is most probably related to the size of PCR amplification product – being too large to be amplifiable with degraded/damaged archival DNA. The sequencing of PCR product amplified with DNA from fresh/frozen samples and then designing suitable internal primers for generating shorter DNA fragments might be the solution for application of this molecular marker to preserved specimens.

A PCR primer set for the **rhodopsin gene** did not give satisfactory results (Fig. 57 (c) and (d)). The problem is most probably related not only to the size of the PCR product, but also to attempts to amplify single/low copy genes from preserved specimens.

All results presented in this section are preliminary and findings need to be confirmed with additional experiments that involve more fish specimens and arDNA extracts, i.e. including sequencing of PCR products. However, the results of these experiments were sufficient to confirm the findings from the previous experiments - the importance of the DNA

extraction protocol used in conjunction with a particular PCR marker system and the significance of preservation method for the specimens usability in molecular investigations. In other words, the extraction methods (including pre-extraction treatments of preserved tissue) used in molecular studies of preserved tissues need to be chosen carefully, in light of the specific requirements of particular investigations.

## 3.5. Introduction of a new term: "archival DNA (arDNA)"

DNA extracted from tissues of museum and other archival collections possess some characteristics that are mostly not present in DNA extracted from fresh and frozen tissues, for example:

- DNA extracts usually contain a variety of PCR inhibitors that are normally not present in DNA extracts from fresh/frozen tissues (diffusible and non-diffusible inhibitors) (An and Fleming 1991; Diaz-Viloria *et al.* 2005; this study);
- It is rarely possible to extract high-molecular-weight DNA, i.e. DNA is usually degraded and fragmented, especially if the tissue/specimen has been stored in preservative for long periods (Tang 2006; Skage and Schander 2007; this study);
- Contaminant DNA may be present (from human handling of specimens, viruses, bacteria, fungi or any parasites present) (Boyle *et al.* 2004; Wandeler *et al.* 2007);
- DNA extracted from archival tissues is likely to be more susceptible to showing false PCR results than high-molecular-weight DNA, i.e. to produce false negative (because of PCR inhibitors present in DNA extract, damaged targeted DNA sequences and/or insufficient amounts of DNA) and false-positive PCR amplifications (because of micro-contamination with exteriors DNA) (Prendini *et al.* 2002; Boyle *et al.* 2004; Tang 2006; Skage and Schander 2007; this study).

From the inevitably incomplete list above, it is possible to see that there are more specific features related to DNA extracts from archival collections than to DNA extracts from fresh/frozen tissues. In my opinion, there is a need for the introduction of a new term for DNA extracted from chemically preserved tissue. My suggestion for this term will be: ***"archival DNA (arDNA)"***.

The characteristics of archival DNA (arDNA) seem to be very similar to those of ancient DNA (aDNA), but have different causes (see Brown (2001); Paabo *et al.* (2004); Gilbert *et al.* (2005a); Cooper (2006) for aDNA). Ancient DNA has been exposed for long periods to different environmental (and other) effects that cause degradation of DNA, but archival DNA has been exposed primarily to chemical effects that cause different types of

degradation, cross-linking and the presence of impurities, making arDNA very difficult (or in some cases impossible) to use in molecular investigations (France and Kocher 1996; Diaz-Viloria *et al.* 2005; Chakraborty *et al.* 2006; Tang 2006). Ancient DNA has not been exposed to preservatives, or where this is the case (e.g. in mummies) deterioration due to age is probably a more important factor. It would not be justified to use the term "ancient DNA (aDNA)" for the DNA extracted from archival collections and specimens of *post mortem* age up to decadal length, despite similarities in the above-stated features between ancient DNA and DNA extracted from archival collections. All archival and museum collections are not necessarily (and mostly are not) "ancient", but archival DNA possesses characteristics that are clearly different from DNA extracted from fresh and frozen tissues. Introduction of this new term "archival DNA (arDNA)" will help in clarification, and provide more accurate definition of the terms "ancient collections" and "ancient DNA", i.e. how old collections or samples need to be in order to be considered as an ancient collection or specimen – hundreds, thousands, or hundreds of thousands of years. On the other hand, using the term "archival DNA (arDNA)" for DNA extracted from preserved samples makes it perfectly clear that this DNA possesses characteristics different from DNA extracted from fresh/frozen samples, and presents difficulties for molecular work with chemically treated samples of a particular age (the majority of them being a few decades old, or even less).

Considering all this, the appropriateness and introduction of the new term "archival DNA (arDNA)" for DNA extracted from fixed and preserved tissue is justified. There have been attempts by other authors to introduce specific terms for DNA extracted from collection specimens in order to distinguish archival DNA from ancient DNA (for instance, Herniou *et al.* (1998) suggested the term "vintage DNA (vDNA)" for DNA extracted from spirit collections). However, in my opinion, the term "archival (arDNA)" has much broader meaning and it is easier to define exactly the source of DNA from any kind of preserved material just by adding an appropriate prefix (e.g., formalin-arDNA, formalin-ethanol-arDNA, dry-arDNA, ethanol-dry-arDNA, or EFEP-arDNA for DNA extracted from ethanol-fixed, ethanol-preserved organisms, FFEP-arDNA for DNA extracted from formalin-fixed, ethanol-preserved organisms, FFSP-arDNA for DNA extracted from formalin-fixed, Steedman-preserved organisms, or formalin/ethanol-arDNA, formalin/Steedman's-arDNA and so on). Also, in my opinion, it is not justified to use the term "ancient DNA" for DNA extracted from archival specimens (chemically preserved samples) 10 years old (or a few decades old), as some authors have suggested (Rey *et al.* 2004; Web page of the Museum Victoria Australia, http://www.museum.vic.gov.au/scidiscovery/dna/ancient_dna.asp).

Recently, Wandeler *et al.* (2007) produced a list of differences between DNA extracted from natural history collections and ancient samples.

Basically, we have a largely incomplete understanding of the nature of archival DNA (the DNA damage) and specific alterations introduced to nucleic acid by different chemicals used for preservation of organisms/tissue, as well as the possibility of reversing some DNA damage in order to make arDNA more suitable for PCR and molecular investigations (Barker *et al.* 2005b; Tang 2006; Gilbert *et al.* 2007b,c,d; Skage and Schander 2007; Wandeler *et al.* 2007). Besides the age/exposure inappropriateness to use the terms "ancient" (very old remains of organisms that were exposed primarily to the natural environmental conditions) and "archival" (chemically preserved specimens which were collected, preserved and stored under controlled conditions – the majority of them just few years/decades old) as synonymous, we do not know the true extent of similarities between archival DNA and ancient DNA regarding the extent of DNA fragmentation, types of crosslinking (covalent and non-covalent, i.e. reversible and non-reversible crosslinks) and DNA modifications. Are the properties of arDNA and aDNA really the same?

# Chapter 4. GENERAL DISCUSSION, CONCLUSIONS AND FURTHER DIRECTIONS

This study has demonstrated that, despite degradation, it was possible to extract PCR amplifiable DNA from museum, fluid-preserved fish specimens of *Nezumia* and mackerel. Molecular investigations on chemically preserved specimens (especially formalin-fixed ones) require a special approach and the development of different strategies in order to overcome problems and enable DNA studies (Tang 2006; Gilbert *et al.* 2007b,c,d; Skage and Schander 2007; Wandeler *et al.* 2007). This study was focused on developing and optimising DNA extraction protocols considering factors that might be important in producing PCR amplifiable formalin-arDNA, as well as developing effective approaches in studying DNA from preserved specimens on species with unstudied genomes.

The limitations of using these preserved specimens for the full potential of DNA analysis resulted mainly from either low arDNA quantity and quality and/or poor DNA extract quality (the presence of PCR inhibitors). Molecular studies on genetically uncharacterised taxa and isolation of markers and DNA sequences from unstudied genomes based exclusively on investigations of archival specimens is complex, but not impossible (see also Boyle *et al.* 2004). This study has demonstrated the feasibility of molecular studies on genetically uncharacterised fish species (*Nezumia aequalis* and *N. micronychodon*) that had been fixed in 10% unbuffered formalin and stored for up to two decades in the Steedman's solution in the Natural History Museum (London, U.K.). Formalin and Steedman fixation/preservation should be avoided whenever possible, but the preservation of existing collections cannot be changed now and it is up to researchers to try to extract PCR amplifiable DNA. Supplementary information from ten differently preserved specimens of mackerel is supporting evidence for correlations between preservation methods of archival specimens, DNA extraction protocols and the possibility of recovering PCR amplifiable arDNA.

There were several reasons why experiments in this study were performed first on formalin-fixed, Steedman's preserved specimens, and then results validated on "good tissue samples" of the *Nezumia* species:

1) Doing the research in this order, it was possible to investigate and evaluate the real possibility of obtaining reliable and reproducible results using specimens exclusively from the natural history collection and species with unstudied genomes. This might potentially be crucial for molecular investigations on extinct and rare species because results obtained from the archival specimens of such species cannot be re-checked and validated against molecular results obtained from fresh/frozen and/or short-period ethanol-stored tissue samples of these organisms.

2) The possibility of cross-contamination between formalin-Steedman-arDNA and "good" DNA from control *Nezumia* samples was eliminated.

At the beginning of this study (i.e. after facing the extreme difficulty of extracting PCR amplifiable arDNA, especially for use in mitochondrial PCRs) attempts were made to obtain fresh/frozen or ethanol preserved samples of *Nezumia* tissue from the UK or any other collection around the world but, at that time, these attempts were without success in securing a reliable positive control for the study. Institutions that were asked did not have these species in their collections or, if they did, they could not guarantee the accurate identification of species, or even genus. The help of a specialist in this group of fish is often needed in order to recognise distinguishing morphological taxonomic characters for accurate and reliable species identification (Iwamoto *et al.* 1999). This underlines the need for, and importance of, developing molecular markers for molecular identification and characterisation of rattail fish species in order to make their identification easier and more reliable which will certainly aid in expanding research on these species and their inclusion in broader biological studies.

## *4.1. DNA recovery from preserved specimens*

The application of an appropriate DNA extraction protocol (including the pre-extraction treatment) on preserved tissue samples is important for success in extracting PCR amplifiable archival DNA (Schander and Halanych 2003; Tang 2006; Ferrer *et al.* 2007; Gilbert *et al.* 2007b), but the species/tissue preservation (the preservation effect) is more critical than the DNA extraction method (Greer *et al.* 1991; Linvillle *et al.* 2004; Ferrer *et al.* 2007; Schill 2007) or the PCR cocktail which can be optimised (Schill 2007).

### 4.1.1. DNA extraction methods and pre-extraction treatments of preserved tissue

Important guidelines in using formalin-fixed (and other archival) specimens for molecular investigations are:

(a) To select the most appropriate DNA extraction protocol (or a combination of protocols) that will produce PCR amplifiable DNA. It is much more important to obtain an arDNA extract that is free of PCR inhibitors than very good yields of DNA. The significance of the right selection of a DNA extraction protocol and its optimisation for preserved specimens has been pointed out by many researchers (Herniou *et al.* 1998; Vince *et al.* 1998; Poljak *et al.* 2000; Tang 2006; Gilbert *et al.* 2007b).

(b) To avoid DNA extraction protocols (chemicals and manipulations) that might cause further damage of the DNA and/or reduce its PCR usability by increasing the amount and/or kinds of diffusible PCR inhibitors in arDNA extracts. For example, phenol-chloroform-isoamyl alcohol purification might not be the best procedure for formalin-fixed specimens because of the possibility that this method could increase the amount of PCR inhibitors in arDNA extracts by the presence of phenolic compounds carried over into the final DNA sample (Simon *et al.* 1996; Wilson 1997) and thus reduce the PCR efficiency, or completely inhibit the PCR amplification (for application to some, or to all PCR markers). Also, it might cause a further fragmentation of DNA during the process (Wang et al. 1994; Skage and Schander 2007).

(c) To select a DNA extraction protocol that is most suitable for investigating a particular DNA region, i.e. for application to a particular PCR marker system/genome region. This study indicates that differently based DNA extraction protocols (guanidinium, PCI, silica spin column) might be more adequate for application to specific PCR marker system/primers. Particular chemicals and isolation procedures under appropriate conditions (pH and temperature) are known to be effective for reversing specific DNA-protein crosslinks, i.e. cleavage of crosslinks (methylene bridges) and liberating proteins that are bound to specific DNA sequences/regions/genes (Toth and Biggin 2000; Barker *et al.* 2005b; Yamashita 2007). Thus, some protocols and buffers might enhanced DNA-crosslink reversal properties (Gilbert *et al.* 2007b) and be more applicable to some target DNA sequences/genes.

(d) The choice of targeted sequence is also important for the successful PCR amplifications because of some being less abundant in the genome (Skage and Schander 2007), the possibility that formaldehyde might cause stronger and less reversible cross-linking with

particular target genes/sequences (Karaiskou *et al.* 2007) and/or DNA sequences with base compositions that are more susceptible to other types of damages (strand nicks, breaks, nucleotide alteration; Rait *et al.* 2006). It is also expected that all PCR primers (including different RAPD primers) and molecular markers would not have the same sensitivity and amplification efficiency in molecular work with difficult samples where the number of available DNA molecules/targeted sequences for PCR might be relatively low (Boyle *et al.* 2004; Willerslev *et al.* 2004; Karaiskou *et al.* 2007; Webb *et al.* 2007) and in the presence of PCR inhibitors (Bucklin and Allen 2004; Diaz-Viloria *et al.* 2005; Tang 2006).

The good efficiency of protocol (A) developed in this study is probably due to a combination of chemical solutions applied in pre-extraction, extraction and purification steps that are beneficial for reversing DNA-protein crosslinks (DPCs) and in producing arDNA extracts with relatively low amounts of PCR inhibitors. Substances/chemical solutions (glycine, SDS, EDTA, DTT, high salt) used in this protocol are known to be effective in liberating crosslinked proteins bound to DNA (Orlando *et al.* 1997; Barker *et al.* 2005b; Yamashita 2007). It also seems to be important in which step (and probably in which buffer) a particular substance/chemical was introduced during pre-extraction and DNA extraction procedure. For example, Gilbert *et al.* (2007b) reported a serious decrease in producing amplifiable formalin-arDNA if 25 mM glycine was added to the incubation buffer during the DNA extraction step with proteinase K digestion. However, this and other studies (Kearney and Stuart 2004; Hasbun *et al.* 2005) have demonstrated that the use of 1xGTE pre-washing buffer (containing 100 mM glycine) is beneficial in producing PCR amplifiable formalin-arDNA. The difference in using glycine in the study of Gilbert *et al.* (2007b) and in its use in this and other studies is that Gilbert *et al.* (2007b) applied glycine to the incubation (digestion) buffer while this and other studies included glycine in the pre-extraction buffer (for rinsing preserved tissue at room temperature prior to the DNA extraction). Findings of Gilbert *et al.* (2007b) suggest that glycine might have the adverse effect if included in this extraction step at a temperature of 55°C and/or carried over into the incubation buffer. Furthermore, Gilbert *et al.* (2007b) results support the observation made in this study that interchangeable washing of preserved tissue in 1xGTE buffer and water, and water applied as a last wash in this pre-extraction treatment of preserved tissue, is more effective in producing PCR amplifiable arDNA (with a lower PCR inhibition) than just washing a tissue sample only in 1xGTE (and/or applying 1xGTE as a last wash). This is probably due to the removal of 1xGTE buffer-substances, including glycine, from a preserved tissue sample before subjecting it to the DNA extraction procedure (digestion with proteinase K and

prolonged incubation at 55°C). Certainly, it is important not only which substances are introduced in DNA extraction procedure of preserved specimens, but also under which physical/chemical conditions, and in which order as well.

The "cocktail" of chemicals when mixed (substances present in preserved organisms, chemicals used for washing tissue samples prior to DNA extraction and chemicals applied in a particular DNA protocol) may affect DNA extractability from preserved specimens, DNA integrity and/or PCR usability of these arDNA extracts. These complex interactions are not fully understood and need to be investigated (Tang 2006), but it seems that they are affecting the reversibility of crosslinks by producing/releasing different kinds of derivatives during a DNA extraction procedure, possibly establishing a new equilibrium in these reactions (Rait *et al.* 2006; Gilbert *et al.* 2007b; Skage and Schander 2007). This consequently might affect the quantity and quality of extracted arDNA, including the ability to generate larger sizes of PCR fragments (Sepp *et al.* 1994).

### 4.1.2. Deep-sea organisms and retrieval of arDNA

Further investigations should be also conducted with special concern to deep-sea organisms because of specific adaptations related to metabolic and oxidative processes, production of substances and accumulation of chemicals that might be related to the particular tissue types and/or deep-sea species, i.e. particular constituents of sea-water related to the specific localities and habitats in the sea (Sole *et al.* 2001; Angel 2003; Treberg *et al.* 2003; Speers-Roesch *et al.* 2006). This, in association with different chemical preservation and DNA extraction protocols, is worth investigating for a range of deep-sea groups of organisms and species because of the possibility that these factors might affect "survival" and extractability of arDNA in such organisms, and might even have different effects on mitochondrial and nuclear arDNA.

For example, it was reported that NaCl is useful in preventing the crosslinking between the macromolecules if DNA is exposed to formaldehyde (Orlando *et al.* 1997; Gebicki and Gebicki 1999; Toth and Biggin 2000; Barker *et al.* 2005b; Yamashita 2007) which is of significance for marine organisms - related to a possible reduction of biomolecular cross-linking in those organisms during formalin fixation. Vachot and Monnerot (1996) have already reported that marine samples fixed in 10% sea-water formalin exhibit better DNA preservation than salt-free unbuffered formalin, but not as good as buffered formalin

solutions. However, we know little about the possible effects of other compounds (such as cadmium, manganese, mercury, lead, cobalt, arsenic, ammonium, methane) that might be present in some deep-sea species and/or in particular tissues (Lilley *et al.* 1993; Fisher 1995; Hannington *et al.* 1995; Angel 2003) and DNA preservation (see Barker *et al.* 2005b).

### 4.1.3. PCR inhibitors

The issue about the PCR inhibition is complicated because an inhibitor may act in more than one way and the relationships between chemical, enzymic and physical factors often cannot be distinguished (Wilson 1997). For example, sodium chloride (NaCl) and Tris are useful in preventing the crosslinking between the macromolecules if DNA is exposed to formaldehyde and peroxidized proteins (Gebicki and Gebicki 1999; Toth and Biggin 2000), SDS and EDTA are useful in reversing crosslinks from formalin-fixed tissues at elevated temperatures (Yamashita 2007), but all these chemicals are also strong PCR inhibitors if they are not sufficiently removed from the arDNA extract during purification steps (Qiagen 2002; McNevin *et al.* 2005). Also, some compounds could be both – inhibitors and facilitators making the matter even more complicated (see Wilson 1997). The application of a DNA extraction protocol and purification steps that are efficient in the removal of PCR inhibitors are often more important for the success of the PCR amplification than achieving a good yield of arDNA (Rossen *et al.* 1992; Simon *et al.* 1996; Wilson 1997; von Wurmb-Schwark *et al.* 2004), although the arDNA yield could be important for generating larger amplicon sizes (Sepp *et al.* 1994) for some target sequences that are low in copy-number (especially relevant if a size of the tissue sample used for DNA extraction is small) (Sepp *et al.* 1994; Karaiskou *et al.* 2007) and for the microsatellite genotyping with longer microsatellites (Watts *et al.* 2007).

#### *4.1.3.1. "Spiking" and diluting arDNA extracts*

Detection of diffusible PCR inhibitors (substances/chemicals that co-purify with arDNA) is very important in order to distinguish them from non-diffusible PCR inhibitors (i.e. the actual damage of DNA induced by preservation chemicals and/or post mortem changes). Because of often not knowing which diffusible PCR inhibitors are present, it is difficult to guess which type of inhibitors they are, and then accordingly to apply an appropriate procedure to eliminate them (Vince *et al.* 1998; Tang 2006). There are different strategies for detecting PCR inhibitors and overcoming the problems of PCR inhibition (Rossen *et al.* 1992; Wilson 1997; Vince *et al.* 1998; Latham 2003; Tang 2006; King *et al.* 2008). The

simplest one is to dilute an arDNA sample (Vince *et al.* 1998), but this is not always a reliable approach, especially if arDNA extracts contain very strong PCR inhibitors and/or too high concentrations of PCR inhibitors in arDNA extracts - this was the case in this study with the majority of arDNA extracts produced by unmodified phenol-protocol (C). "Spiking" a DNA sample (mixing arDNA and "good" DNA) is a much more reliable way to establish the presence of diffusible PCR inhibitors in the DNA extracts (Tuross 1994; Arroyo-Pardo *et al.* 2002; Gilbert 2003a; Pusch and Bachmann 2004; Serre *et al.* 2004; von Wurmb-Schwark *et al.* 2004; Willerslev *et al.* 2004). In this study, control DNA from fresh/frozen cod and rainbow trout was mixed with formalin-arDNA extracts of *Nezumia* and successfully applied in detecting "diffusible" inhibitors. These experiments undoubtedly confirmed the presence of "diffusible" PCR inhibitors in some of formalin-arDNA extracts and their connection to a particular DNA extraction protocol (arDNA extracts produced by phenol based DNA extraction protocols showed a stronger PCR inhibition than guanidinium based protocols or silica based commercial kits).

#### *4.1.3.2. Strategies to overcoming PCR inhibition and improving the PCR efficiency*

Simply diluting the DNA extract is an easy, and in many cases successful, approach to overcoming PCR inhibition (Vince *et al.* 1998; Boman *et al.* 1999; Kalmar *et al.* 2000; Olson *et al.* 2005), but excessive dilution may also dilute the DNA to non-amplifiable concentrations (Wilson 1997; Mulligan 2005; King *et al.* 2008), as might be the case with some dilution experiments in this study. The DNA extraction method, including pre-extraction treatments of preserved tissue and the method of precipitating and purifying DNA, is important for DNA recovery and removal of PCR inhibitors from DNA extracts (Rossen *et al.* 1992; Simon *et al.* 1996; Wilson 1997; Radstrom *et al.* 2004), but optimisation of the DNA amplification conditions by the use of alternative DNA polymerases (inhibitor-tolerant thermostable polymerases) and/or amplification facilitators also can decrease (eliminate) the PCR inhibition effect and increase the PCR efficiency (Simon *et al.* 1996; Wilson 1997; Radstrom *et al.* 2004; Hoorfar *et al.* 2004; Valasek and Repa 2005; King *et al.* 2008).

Isopropanol precipitation in comparison to ethanol precipitation results in higher yields of DNA and better removal of PCR inhibitors from DNA extracts (Hanni *et al.* 1995; Montiel *et al.* 1997; Rohland *et al.* 2004; Mulligan *et al.* 2005). The addition of polyacrylamide for DNA precipitation seems to be beneficial for removing PCR inhibitors and producing PCR amplifiable arDNA extracts. An additional purification of DNA samples, such as: Dextran

Blue (Kalmar *et al.* 2000), dialysis of DNA extracts (Kiesslich *et al.* 2002), silica-based methods (Bouwman and Brown 2002), "Sigma" GenElute PCR purification Kit (Brzuzan *et al.* 2004) etc. might be potentially useful approaches for reducing (eliminating) PCR inhibitors from DNA extracts (this needs to be assessed in further investigations with arDNA). Columns from commercially available DNA extraction kits (e.g. Qiagen or Promega) and use of microconcentrators (e.g. Vivaspin filter-concentrators) proved to be useful as additional purification steps in this study. It is important to select purification procedures that are effective in eliminating inhibitors but not eliminating DNA from arDNA extracts and/or introducing DNA contaminants (Mulligan *et al.* 2005).

An adequate choice of thermostable polymerases and "hot-start" PCR (with polymerase-antibody complexes rather than application of wax beads) can also be used to overcome PCR inhibition and/or to increase the yield of a PCR product (Wilson 1997; Boyle *et al.* 2004; Quach *et al.* 2004; Bhadury *et al.* 2005; Valasek and Repa 2005). PCR facilitators such as bovine serum albumin (BSA), glycerol, dimethyl sulfoxide (DMSO), Tween 20, etc. may resist PCR inhibitors and improve amplification of DNA (Simon *et al.* 1996; Wilson 1997; Radstrom *et al.* 2004; Hoorfar *et al.* 2004; King *et al.* 2008). However, tests with some of these reagents did not show significant improvements of PCR amplifications in this study, except, perhaps, the use of the Q-Solution (containing glycerol) from the Qiagen PCR Kit for some mitochondrial PCR amplifications.

A few simple, but useful, laboratory-technical approaches for increasing the chances of amplifying archival DNA, adopted in this study and also presented by others in molecular work with degraded and damaged DNA (e.g., Legrand *et al.* 2002; Mulligan 2005; Morin *et al.* 2006), are: using increased amounts of *Taq* polymerase (usually 1-2 units per PCR reaction), applying a lower annealing temperature and a higher number of PCR cycles (usually 40-50). Some researchers claim that the application of a higher concentration of DNA template is also beneficial to better amplification of degraded DNA (e.g. Sefc *et al.* (2003) for microsatellite amplification from museum feather samples), but the results from this study do not fully support this, mainly because of the presence of PCR inhibitors that significantly inhibit PCR if larger volumes of arDNA extract are used. Similar observations to those in this study were also reported by Rossen *et al.* (1992) for food samples, Ciesielski *et al.* (2002) for aDNA, or von Wurmb-Schwark *et al.* (2004) for forensic samples.

Nested and touch-down PCRs might be advantageous for overcoming the PCR inhibition and increasing the efficiency and/or specificity of PCR amplifications with DNA from

preserved and other difficult samples (Simon *et al.* 1996; Jovanovic *et al.* 2003). Touch-down PCR was tested for a few RAPD-PCR experiments in this study, but the number of experiments was insufficient to draw a clear conclusion as to whether this approach improved RAPD-PCR amplifications.

As well as the use of optimised DNA extraction protocol steps and PCR conditions, pre-extraction preparation of the preserved tissue (pre-washing and/or drying) is also important in order to remove residuals of formalin and other chemicals/substances from a tissue sample that might act as PCR inhibitors (Shiozawa *et al.* 1992; Shedlock *et al.* 1997; Wirgin *et al.* 1997; Li *et al.* 2000; Fang *et al.* 2002; Shi *et al.* 2002, 2004; Klanten *et al.* 2003; Schander and Halanych 2003; Kearney and Stuart 2004; Hasbun *et al.* 2005; Austin and Melville 2006; Tang 2006; Gilbert *et al.* 2007b; Karaiskou *et al.* 2007). In this study interchangeable washing of a preserved tissue sample in 1xGTE buffer (Shedlock *et al.* 1997) and water, then freezing and drying of a tissue sample at room temperature prior to DNA extraction proved to be good pre-extraction treatments for reducing the amounts of formalin and other substances from a preserved tissue sample.

The application of the real-time PCR technology should be applied in work with archival DNA because it can help in clarifying this issue of the quality/quantity of arDNA and inhibitor concentrations in arDNA extracts (Tang 2006; Gilbert *et al.* 2007b,c,d).

### 4.1.4. Smearing tests/assessments

PCR products viewed in this study as "smearing" (or only smears instead of visible PCR-DNA bands) indicated a poor quality (and/or a low quantity) of applied PCR-template. Similar observations were reported by other researchers who applied formalin-arDNA as PCR-templates (Chakraborty *et al.* (2006) on formalin-arDNA of fish; Bhadury *et al.* (2006a) on formalin-arDNA of nematodes). Amplified DNA which appear as smears on a gel instead of distinct bands probably can be used as a rapid test and a fast assessment of whether extracted arDNA from a particular specimen is usable for PCR or not. This kind of test should be performed by applying different *Taq* polymerases and PCR markers (primers) because of possible differences in PCR efficiencies with different *Taq* polymerases and primers (Boyle *et al.* 2004; Willerslev *et al.* 2004; Bhadury *et al.* 2005, 2006a,b; Valasek and Repa 2005).

### 4.1.5. Sizes of PCR amplicons generated with formalin arDNA

Many published papers on formalin-fixed tissue have reported the inability (or extreme difficulty) of generating PCR-DNA fragments larger than 300-400 bp (e.g. Shibata 1994; Bonin *et al.* 2003; Liborio *et al.* 2005; Stanta *et al.* 2006; Karaiskou *et al.* 2007). However, amplification results from this study suggest that formalin-arDNA is not as fragmented as previously thought. For example, RAPD fragments with sizes often larger than 800 bp (or even 2000 bp) were generated in this study without applying any special treatment to repair damaged DNA. Some other researchers also found large DNA fragments from formalin-fixed tissue (e.g. Goelz *et al.* (1985) detected DNA fragments up to 10,000 bp by using restriction endonucleases; Savioz *et al.* (1997) successfully amplified an 838 bp long fragment from 46 year old preserved tissue; Fang *et al.* (2002) amplified DNA fragments up to 2000 bp from different vertebrate tissues stored for 16-70 yrs in unbuffered formalin; Inadome and Noguchi (2003) detected a large arDNA fragments from formalin-fixed tissue samples by applying HPLC analysis; Wan *et al.* (2006) found fragments larger than 2 kb in formalin-fixed faeces samples stored for a couple of years).

The size of DNA fragment for application to a particular marker might not be always of crucial significance for successful PCR amplifications (see Skage and Schander 2007), and/or be a good indicator of arDNA fragmentation - the possibility that particular targeted gene regions/ DNA sequences are more prone to crosslinking (Karaiskou *et al.* 2007), or for other reasons (e.g. the use of particular DNA extraction protocol in producing arDNA extract, thermostable polymerase, or primer matching; Boyle *et al.* 2004; Bhadury *et al.* 2005, 2006a; Skage and Schander 2007). However, in general, it is expected that smaller fragment sizes are easier to amplify with a higher degree of reproducibility and consistency than fragments of larger sizes (O`Leary *et al.* 1994; Boyle *et al.* 2004; Miething *et al.* 2006; Gilbert *et al.* 2007b,c) because of degradation and DNA damage that certainly occur during fixation and preservation storage (Schander and Halanych 2003; Tang 2006; Gilbert *et al.* 2007b,c,d; Skage and Schander 2007; Wandeler *et al.* 2007).

The data from this study strongly suggest that the number and sizes of RAPD-PCR bands (fragments) generated with particular RAPD primers can indicate a degree of degradation and fragmentation of arDNA. In other words, if it is possible to generate a higher number of RAPD-PCR fragments and if among these fragments are also RAPD-PCR fragments of larger sizes, this might indicate a lower level of DNA fragmentation caused by preservation (and/or by death of an organism). However, the size and number of bands are often related to

the use of a particular RAPD primer, PCR-DNA template concentration (i.e. dilution of arDNA), the presence of diffusible PCR inhibitors in a particular arDNA extract, DNA extraction protocol used and pre-extraction treatments (washing/drying of tissue samples), as well as the type of *Taq* polymerase and PCR conditions. Because of this, it is recommended that RAPD-PCR experiments are performed with a few different RAPD primers and a range of arDNA extracts produced by different extraction protocols, with dilutions, and with different *Taq* polymerases in order to assess accurately the quality of arDNA and arDNA extracts. The RAPD assay can be used as a rapid, reliable and cost-effective way of assessing the quality and usability of extracted arDNA. Other researchers also found RAPD-PCR methodology a good approach for investigating the quality of extracted DNA (Adams *et al.* 1999; Carvalho and Vieira 2000; Siwoski *et al.* 2002; Iniguez *et al.* 2003; Tayutivutikul *et al.* 2003).

### 4.1.6. DNA sequences derived from formalin-arDNA

There are many controversial results and opinions related to the DNA damage and sequence alterations induced by formalin and the reliability of sequences extracted from formalin preserved material (see section 1.4.1.1.2.1. in Chapter: General introduction; p 46). A consistency and fidelity of sequencing results in this study supports the possibility of obtaining accurate and reproducible sequences from formalin-fixed fish specimens. This finding is related to the 458 bp-long, well-conserved and anonymous genome region of *Nezumia*, but it is not known if this kind of result is valid to any targeted sequences, or only to particular genome regions that are less prone to the DNA damage and to "strong" (non-reversible) crosslinking.

There is an urgent need to establish what factors influence the fact that some sequences are less/more susceptible to nucleotide alterations and being more/less reliable than others. Is it due to the fixation/preservation method: buffered, non-buffered, sea-diluted buffered and unbuffered formalin in conjunction with age and storage conditions of collections; post-mortem effect, i.e. the collection and fixation-delay effect; size of tissue used for DNA extraction; DNA extraction method; targeted DNA sequences; or something else? The priority is to establish what the state of preservation of DNA is in the presence of formalin (types and abundance of particular DNA damages) and to characterise the physical and chemical states of crosslinks. There are suggestions that capillary electrophoresis and mass spectrometric techniques could be used to assess the integrity of DNA and observe damage

in formalin-arDNA (Tang 2006). Carter (2003) adapted denaturing gradient gel electrophoresis (DGGE) and heteroduplex analysis for the estimation of DNA damage in preserved specimens, since these methods allow amplicons of the same length to be separated on the basis differences in their sequence. These methods can detect small changes in DNA sequences – single base substitutions, deletions, insertions and crosslinks (Carter 2003; Schlotterer 2004). Single cell gel electrophoresis (SCGE) also proved to be useful in detecting crosslinks induced by formalin (Liu *et al.* 2006). All these techniques and methods have the potential of revealing damage present in archival DNA and thus facilitate in accruing knowledge about the nature of archival DNA and the accuracy of sequence information extracted from preserved material.

There is also a need to establish if "hotspot damage" is also relevant to preserved material as was found for ancient DNA (Gilbert 2003a,b; Gilbert *et al.* 2005b, 2007a; Binladen *et al.* 2006) and some other samples related to heating of DNA (Banerjee and Brown 2004) and laboratory techniques used (Brandstatter *et al.* 2005). The A/G mutations found in this study are considered to be a genuine mutation (probably SNP) because repeated sequencing (different arDNA extracts of one fish individual) showed always the same base at the same position in the DNA sequences. Differences were only between different fish individuals. Sikora *et al.* (2006) reported reliable analysis of the genuine G/A transitions (the substitution of codon GAC for AAC) for establishing diagnosis of Creutzfeldt-Jackob disease by using formalin-fixed, paraffin-embedded tissues. This indicates that it is possible to extract reliable and reproducible sequences from formalin-fixed material which will allow sensitive analysis, such as detecting the genuine base substitutions or polymorphisms. Recently, single nucleotide polymorphisms (SNPs) have been put forward as suitable for genotyping of degraded DNA because it requires only short target DNA sequences (Thompson *et al.* 2005; Sanchez and Endicott 2006; Morin *et al.* 2007).

However, many issues need much more clarification and understanding of the formalin-arDNA. For example, Williams *et al.* (1999) and Quach *et al.* (2004) found a correlation between artifactual mutation in formalin-fixed material and the number of cells used in PCR – the fewer cells, the more artifacts. They also found that A-T transition mutations (DNA-damage-induced errors) were more prevalent in fixed than in fresh tissue, whereas others expect that G-A (i.e. C-T) transition should be a more frequent type of erroneous sequence substitution due to the deamination of cytosine residues (Wandeler *et al.* 2007) as it was indeed found in DNA extracted from dry museum specimens (Sefc *et al.* 2007) and in ancient DNA (Cooper 2006).

PCR cloning is one of the approaches to investigate the authenticity and reproducibility of DNA sequences extracted from "difficult" material (Gilbert 2003a; Paabo *et al.* 2004; Wandeler *et al.* 2007). During this study, one PCR/cloning artifact was recorded (repetition of the reverse complementary RAPD primer as a cloned insert), but the artifact was obvious and easy to identify. However, the usefulness of this approach might be less if the existence of "hotspot" damage in preserved material is discovered, i.e. that DNA-damage-induced errors might happen in each cell at the same place in a DNA sequence (the systematically biased alterations; Tang 2006). If fresh/frozen or short-term ethanol preserved tissue sample is available for particular species, it is not a problem to reveal such errors, but it is very difficult for extinct, rare and difficult-to-collect species that are availably only as preserved material.

## *4.2. RAPD methodology and archival DNA*

During the project, attempts were made towards optimisation of the RAPD-PCR methodology on investigated *Nezumia* specimens, but RAPDs are erratic and definitely cannot be used with formalin-Steedman archival DNA for their traditional applications (presence/absence of bands and RAPD-distance analysis). Besides the inconsistency of RAPD-PCR fragments (bands) in profiles (see also Eckerman and Welsh 1997; Siwoski *et al.* 2002), the intensity of co-migrating bands was often different, even in PCR duplicates. Variation in the intensity of amplicon bands was also found by other researchers when DNA from formalin-fixed tissues was used (James *et al.* 2002). Such a great inconsistency of RAPD results (patterns and RAPDs) with formalin-arDNA is probably related to the heterogeneity of DNA samples, i.e. damage to the DNA and to the presence of PCR inhibitors in arDNA extracts. Even 95% ethanol-preserved specimens stored at room temperature for 1 year were reported as unsuitable for traditional RAPD analyses because of generating RAPD profiles with missing bands (Carvalho and Vieira 2000).

However, in this study, the RAPD methodology was proved to be invaluable as a basic molecular tool in gathering necessary information on formalin-arDNA extracts (testing the efficiency of DNA extractions/amplifications) and for developing sequence-specific primers. Initially, at the beginning of this project, RAPD methodology was almost the only choice of molecular technique to begin any molecular investigation on museum formalin-fixed specimens of *Nezumia*. At the time, it was impossible to know whether PCR amplifications

tested with specific primers (for mitochondrial targeted gene sequences) were failing because of the inappropriateness of primers that were used in experiments (not matching the sequence of *Nezumia*), or if it was a problem related to extracted arDNA (PCR non-amplifiable arDNA) and DNA extraction protocols. The situation was even more complex because a reliable and effective method for measurement of arDNA concentration was not immediately found (because of arDNA degradation - it was not possible to apply gel molecular weight marker measurement of DNA concentration; low concentrations of extracted formalin-arDNA and impurities that co-purify with DNA made the measurement of archival DNA by the spectrophotometer unreliable and difficult). All this made it very difficult to know anything for certain about the molecular markers applied, formalin-arDNA extracts and the usefulness of the DNA extraction protocols applied. The only information was that most PCR experiments were failing to yield any PCR products, except with some of the tested RAPD primers.

Because of the lack of sequence information on *Nezumia*, inapplicability of tested "universal" primers to preserved specimens, unavailability of fresh/frozen (or more suitable preserved tissue samples) of *Nezumia* that could be used for testing these "universal" primers and for developing molecular markers, and little knowledge of the nature of archival DNA (i.e. damage to DNA and its survival in preserved specimens), only RAPD-PCR methodology proved to be a useful molecular approach in this study. The use of RAPDs as a basis for further investigation and development of sequence-specific markers (i.e. STSs and SCARs) and other classes of molecular markers is a feasible and important strategy in studying genetically uncharacterised species that are available mostly (or only) as preserved specimens (especially if available "ready" primers are not applicable to the investigated specimens).

Designing the sequence-specific, RAPD-derived primers that will amplify different sizes of DNA fragments will secure their applicability to specimens with degraded DNA, because of a higher probability of generating amplicons of a smaller size from such samples (O'Leary *et al.* 1994; Chase *et al.* 1998a; Boyle *et al.* 2004; Miething *et al.* 2006; Gilbert *et al.* 2007b,c), and because of the different amplification efficiency of particular primers. By generating many such primers, it is possible to apply a multiplex PCR for a fast screening of arDNA extracts, specimens, populations, species, and so on. In other words, if a multiplex PCR contains primers that amplify targeted sequences of different sizes for a particular species, only one PCR reaction can facilitate fast screening of the quality of arDNA extracts produced by particular DNA extraction protocols. Developed primers and multiplex PCRs

might also be applied for the fast screening of differences between populations, species and other taxa if they proved to be specific for a particular rank. Multiplex assays on preserved specimens proved to be a useful approach for assessing arDNA extraction and obtaining molecular information from archival collections, but currently they are applied only for species and/or human tissues for which primers are already developed (Gilbert *et al.* 2007b,c). The RAPD approach gives the opportunity for multiplex assays to be developed and applied on non-model species and species with unstudied genomes that are available only/mostly as preserved specimens and for which "ready" available primers are not applicable for different reasons.

General suggestions that RAPDs usually contain repetitive DNA (Ghany and Zaki 2003; Yang *et al.* 2006; Eckerman 2006), that RAPDs are useful in identifying conserved regions of genomes (Mehling *et al.* 1995; Mellado *et al.* 2000), that RAPDs may contain regions of retrotransposons, i.e. transposable elements (Dioh *et al.* 2000; Abe *et al.* 2005), and that some of retrotransposable elements may contain parts of protein coding sequences (Banki *et al.* 1994; Mourier 2005) provide great potentials for molecular investigations and are of interest for investigations related to the application of RAPD-PCR methodology to preserved organisms. The revealed 458 bp region of *Nezumia* contain an ORF and a possible SNP providing a significant starting point for more complex molecular investigations for *Nezumia* species. Also, BLAST searches indicated that a well-conserved genome region of 458 bp (A-T rich region with ORF), a sequence that occurs in both investigated fish species of *Nezumia*, might be a part of the gene responsible for a membrane protein of the Rhomboid family. This needs to be substantiated by further investigations, but the results from this study emphasise how useful RAPD-PCR methodology can be from different aspects for application to preserved specimens, especially for species with unstudied genomes.

Furthermore, sequencing of RAPD-fragments might be one way towards genomics since RAPD offers the opportunity to access information about a large portion of the genome, and by sequencing many RAPDs, it is possible to reveal a significant amount of sequence from different species (Liu *et al.* 2006), including taxa with unstudied genomes and specimens that are available only as preserved material.

## *4.3. Recommendations in molecular work with archival DNA*

Based on the experience from this study, the recommended approach for the molecular study of preserved archival specimens will be:

1. Apply the DNA extraction protocol that produces the most successful arDNA extracts, based on information from previous research with a particular preservation method and species. Protocol (A) developed in this study should be applicable to the majority of preserved specimens (with or without modification).
2. Apply RAPD-PCR methodology (with at least three different PCR-DNA template concentrations in duplicates) in order to test the efficiency of DNA extraction/amplifications. Assessment of arDNA extracts should include at least one different molecular marker (e.g. mitochondrial) because of the possibility that some DNA extraction protocols might be more suitable for a particular PCR marker system.
3. If one RAPD primer does not generate PCR products, test a particular arDNA extract with a few other RAPD primers.
4. If PCR products are not generated with RAPD and other primers, test for the presence of inhibitors by:
    a. diluting arDNA extracts and
    b. mixing arDNA with control, "good" DNA ("spiking" arDNA).
5. If "diffusible" PCR inhibitors are detected, an additional purification of arDNA extracts is required.
6. If "diffusible" PCR inhibitors are not detected, but RAPD-PCRs and PCRs with other molecular markers are still unsuccessful, try to apply a different "type" of DNA extraction protocol.
7. If a few different DNA extraction protocols did not produce arDNA extract that is PCR amplifiable with RAPD and other molecular marker systems, and if "diffusible" PCR inhibitors are not detected, it is likely that DNA in the preserved specimen is so damaged/degraded that the specimen is not useful for molecular investigation.
8. If arDNA is extractible and amplifiable, DNA extraction protocol and PCR conditions need to be optimised for preserved specimens under investigation in order to obtain the highest possible level of extraction/amplification reproducibility. After optimisation and selection of robust RAPD primers, convert RAPDs into STS/SCAR markers for more specific analyses that will cover larger portions of the genome and apply other appropriate molecular markers (mitochondrial, ribosomal,

microsatellites, SNPs), if possible, for DNA investigation.

9. All precautions and criteria for working with difficult samples should be respected in order to obtain reliable data. Validation of results in another laboratory is recommended, or at least validation on specimens from another collection. If possible, validation of results on "good" samples (fresh/frozen, short-term ethanol preserved) for extant species will be beneficial for this kind of investigation. However, if it is not possible to arrange this kind of data validation, reproducible results in duplicate samples (duplicate extractions and analyses from the same specimen/collection) may be acceptable for testing the authenticity and validity of the results.
10. Designated areas for DNA extractions, PCR experiments and post-PCR analyses should be organised in the laboratory for work with archival DNA (Bhadury *et al.* 2006a; Wandeler *et al.* 2007). Ideally, areas for working with archival and "good" DNA should be separated (at least for performing DNA extractions and PCR set-ups). Basically, all criteria and standards that apply in work with ancient DNA (Brown *et al.* 1998; Cooper and Poinar 2000; Poinar 2003) and in forensics (Buckleton *et al.* 2005) should be also applied to work with archival DNA and preserved specimens.

Because of possible difficulties in testing the applicability of universal primers on specimens preserved for long periods (especially formalin-fixed ones), these tests and the development of molecular markers for genetically uncharacterised species are easier to do first on good sources of DNA (fresh/frozen or short-period ethanol preserved specimens) if possible, and then perform testing and molecular analyses on archival specimens. RAPD profiling and converting RAPDs into STS/SCAR markers are faster and more reliable on good sources of DNA, but this does not mean that all developed markers on "good DNA" will be applicable to preserved specimens (due to fragmentation and arDNA damage). If experiments are performed in this order, extra precautions need to be undertaken due to possible cross-contamination between control and archival DNA. It is recommended that all experiments with control specimens are performed in a separate room from the room with the archival experiments. It will be even better if experiments with control and archival specimens are performed in different buildings/laboratories because of a high-risk of contamination with "the same" DNA. Beside STS/SCAR primers, it is recommended that mitochondrial, ribosomal, microsatellites and other marker primers are tested on "good" DNA and then developed internal primers for amplifying a smaller size of fragments (preferably species-specific primers) (Chase *et al.* 1998a; Boyle *et al.* 2004).

If good sources of DNA are not available for investigated species with unstudied genomes and if developed primers for specific molecular markers (universal, specific for particular species and/or group of organisms, or primers that were designed from available sequences of other species) are not applicable to the investigated preserved specimens, recommended approaches presented in this thesis are probably the best strategy for generating molecular data. Sequencing of both DNA strands of a few clones and/or PCR products (from the same individual and from different arDNA extracts) should always be performed with archival DNA due to possible misincorporations of bases, or other sequence alterations that might be induced by chemical preservation and during PCR (Tang 2006).

### 4.3.1. Further directions in archival DNA research

Beside the optimisation of DNA extraction method, there are some alternative methods and approaches that might ease the use of arDNA and enable access to more genetic information from preserved specimens of archival collections (Tang 2006; Skage and Schander 2007; Wandeler *et al.* 2007).

#### *4.3.1.1. Possible further optimisations of protocol (A) developed in this study*

Although protocol (A) exhibited a good performance in this study, it is probably possible to make some further improvements and optimisations of the protocol in order to be applicable to the variety of preserved specimens and molecular markers/assays. For example, it would be worthwhile testing: combining some of the extraction steps from protocol (A) and application of silica, Chelex resin, commercially available kits, and/or testing pre-extraction treatments such as a gradual dehydration of preserved tissue with ethanol (Fang *et al.* 2002) – alone and in combination with 1xGTE buffer, testing whether heating the tissue (>80°C) under the influence of pH higher than 7 prior to DNA extraction (hot-alkali treatment; Shi *et al.* 2002, 2004) can produce a better DNA yield and arDNA extracts with less PCR inhibitors. It is known that sometimes even a slight variation in a protocol can make a big difference in the success of DNA extraction of preserved tissue (Tang 2006).

If possible, modification of protocol (A) should also be directed towards simplification: the use of a smaller numbers of reagents to minimize the possibility of foreign DNA introduction through reagents and sample manipulations, as well as shortening extraction time. Partial extraction/amplification success with protocol ($G_3$) tested in this study and

simplified protocol of Cawkwell and Quirke (2000) might be a new direction in simplifying protocols for extraction of PCR-amplifiable archival DNA.

Because of the sensitivity in extracting PCR amplifiable arDNA, it is probably unrealistic to expect the development of a very robust DNA extraction protocol that will work perfectly with all types of preserved specimens and/or produce arDNA extracts that are equally successful in all PCRs (see also Tang 2006). Small differences in tissue samples (size, weight, differential fixation/preservation of specimens) and the number of tissue cells involved in the preparation of arDNA extracts might influence the behaviour of arDNA extracts (Williams *et al.* 1999; Quach *et al.* 2004). Some of these parameters are impossible to control precisely. It will help if molecular biologists are supported by researchers with expertise in chemistry, biochemistry, and biophysics because it will be easier to identify which chemicals and substances cause the inconsistency of arDNA extractions and PCR amplifications. If common PCR inhibitors (diffusible and non-diffusible) are identified, molecular work on preserved specimens will be faster and more successful, although variations between collections are expected because of non-standardised fixation and preservation procedures of samples (Schander and Halanych 2003; Tang 2006; Wandeler *et al.* 2007). Currently, the mechanism of PCR inhibitors, the chemistry of DNA degradation, the state of DNA preservation and the DNA damage of preserved specimens is largely unknown (Vince *et al.* 1998; Schander and Halanych 2003; Tang 2006). All these need much better characterisation in order to apply effective steps for an adequate optimisation of a DNA extraction method.

A particular emphasis should be given to archival mtDNA because of the controversy of the published results and some specificity that might occur with mtDNA (see section 3.3.2.; p 217). In this study, mitochondrial PCR amplifications with formalin/Steedman-arDNA of *Nezumia* proved to be extremely difficult. There is no clear explanation at the moment, but this might be caused by some kind of cross-linking that does not allow sufficient release of DNA from mitochondria with protocols applied in this study ("trapped DNA"; Kiernan 2000; Yamashita 2007) or because of the severe damage of the mitochondrial genome (mtDNA), but most probably the combination of both. These difficulties with the extraction/amplification of mtDNA are found not only to be related to formalin-fixed specimens, but also to museum dry collections (Gilbert et al. 2007d). The state of mtDNA is particularly important to clarify because of a general belief that it is much easier to amplify regions of mtDNA than nDNA from formalin-preserved and other difficult samples because of many-copy numbers of mitochondrial genomes per cell. However, there are opinions that

formalin-mtDNA is less accessible than is nDNA (Tang 2006), or some researchers even question a "survival" of mtDNA in formalin-fixed specimens (Diaz-Viloria *et al.* 2005; Chakraborty *et al.* 2006).

#### *4.3.1.2. DNA repair*

The answer to the recently asked question (Tang 2006; Skage and Schander 2007) if it is more important to continue developments and improvements of DNA extraction protocols, or is it more important to find a way for an effective repair of damaged formalin-arDNA, will probably be - both. The results from this and other studies (see section 4.1.1.; p 233) strongly suggest that optimisation of DNA protocol (and maybe pre-extraction treatments) is important for successful and effective extraction of DNA from preserved specimens. This is important from the aspect of a better arDNA yield (although this aspect is not always of crucial importance for arDNA), removal of PCR inhibitors from arDNA extracts (crucial for extracting amplifiable arDNA), better PCR efficiency and applicability to different PCR marker systems and primers. However, damaged arDNA can to some extent be repaired physically and chemically after extraction from preserved specimens (Bonin *et al.* 2003; Hajibabaei *et al.* 2005; Tang 2006; Skage and Schander 2007).

For example, there were successful attempts by Bonin *et al.* (2003) to amplify longer sequences (up to 300 bp) by applying a pre-PCR restoration treatment (filling single strand breaks followed by a vigorous denaturation step) to DNA extracted from formalin-fixed, paraffin-embedded post-mortem tissues. Bonin *et al.* (2003) associated the use of non-buffered formalin with random single strands breaks and extensive degradation of nucleic acid. The DNA restoration process is based on the ability of the polymerase reaction to restore the nicks after DNA re-hybridisation using the other strand as a template.

A "PreCR" method involves treating damaged DNA *in vitro* with a mixture of DNA repair enzymes before PCR (Tang 2006). The PreCR enzyme mix is supposed to repair abasic sites, nicks, gaps deaminated cytosine and some forms of oxidative damage, but DNA cross-links or highly damaged and fragmented DNA could not be repaired effectively. This method was not successful when tested on formalin-fixed samples (Tang 2006). Restorase, an enzyme mix (Sigma-Aldrich) containing Accu*Taq* DNA polymerase and a DNA repair enzyme, was successful to a degree in increasing the PCR efficiency for amplification of some genome regions (COII and ITS2), but not for the barcoding COI region (Skage and Schander 2007).

From only these few examples, it is noticeable that DNA repair has the use, but also limitations in repair of formalin-arDNA damage. Until we learn more about the state of arDNA extracted from differently preserved specimens and establish some general guidelines regarding the possible DNA expectations from collections of different age and fixation/preservation procedures, it is difficult to improve effectively either the DNA extraction method or repair of the arDNA. An additional, complicating factor to achieve this in full is variation in curatorial processes that might cause significant variations in the chemistry of DNA degradation between collections/specimens which it is claimed are preserved in the same manner (Prendini *et al.* 2002; Tang 2006).

### *4.3.1.3. New methodologies and technologies*

Advances in molecular technologies provide an approach for retrieving a large amount of genetic information using small amounts of tissue/DNA. Some methods that could potentially be used on preserved specimens of natural history collections (Tang 2006) and worth further attention are:

*PCR-mass spectrometry method* (such as matrix-assisted laser desorption/ionization time-of-flight; MALDI-TOF, or sequencing in the mass spectrometer by fragmentation) was suggested as potentially useful for a sequencing/multiplex-genotyping of formalin-preserved specimens (Jaremko *et al.* 2005; Tang 2006).

*Single-molecule sequencing by synthesis* – a method designed for short-read sequencing of genomes (sequencing fingerprints up to 5 bp in length) (Braslavsky *et al.* 2003).

*Whole genome amplification (WGA)* technique requires a small amount of template DNA in order to amplify the genomic DNA (Tang 2006). Most of these protocols rely on high quality template DNA, but Wang *et al.* (2004a) developed a method tolerant to sample degradation.

*Pyrosequencing* (using emulsion polymerase chain reaction) has already been applied in ancient DNA studies (Poinar *et al.* 2006), but not yet assessed on natural history collections.

*Multiplex PCR with minisequencing (MPMS)* is developed by Gilbert (2007c) as a high-throughput SNP typing method for formalin-fixed tissue.

Improved bioinformatics can be used to assemble and analyse short DNA fragments produced by these techniques (Tang 2006; Skage and Schander 2007). These techniques

have great potential for use on natural history collections and can enable access to a large amount of genetic information, but they have some limitations. First, most of these techniques require developed primers for their use, which could be time-consuming and expensive for application to non-model organisms and species with unstudied genomes, and especially difficult for species that are available mostly/only as preserved specimens (extinct, rare and difficult-to-collect species). Another limitation on the use of the majority of the above technologies is that they are extremely expensive (for example, pyrosequencing costing thousand of pounds per run). So, their application is available only to very limited numbers of academic/research institutions with sufficient financial resources.

In comparison to these high-throughput methodologies, the RAPD-PCR methodology is affordable to molecular laboratories with modest facilities. It is not a high-throughput methodology, but the RAPD approach developed in this study is an effective, reliable and low-cost approach to studying preserved specimens from museums and other archival collections.

## *4.4. Handling/ storage/ fixation of newly obtained material*

As already mentioned, it is not possible to change now the preservation of existing archival collections, but for newly obtained material for museums and other collections one should consider preservation and storage that will, besides good morphological preservation, provide good DNA preservation and applicability of the material to DNA investigation. Clearly, the best method for preserving DNA in good condition is crypreservation (at -80°C, or in liquid nitrogen) (Dessauer *et al.* 1996; Prendini *et al.* 2002; Ferrer *et al.* 2007). It is, understandably, unreal to expect that a range of whole specimens will be cryopreserved (especially specimens of big animals), but tissue samples of collected organisms should be. The results from this and other studies strongly suggest avoiding formalin-fixation and Steedman`s fixation/preservation of specimens for DNA investigation. Ethanol preservation is recommended, but not for longer than a few years, unless regularly changed and/or kept at a lower temperature (ideally at 4°C, or at -20°C) (Oliveira *et al.* 2002; Hajibabaei *et al.* 2005; Vink *et al.* 2005). However, these require investigation to establish whether storage of ethanol-preserved specimens at low temperature allows good DNA preservation over long periods and for all species. Isopropanol fixation/preservation did not show good preservation of DNA in investigated preserved specimens of mackerel and this preservation should probably be avoided for fish specimens. However, some authors reported relatively good

DNA results from fish specimens preserved in isopropanol (Shiozawa *et al.* 1992; Wirgin *et al.* 1997), although their samples were previously fixed in formalin. In this study, a control *Nezumia* sample that was preserved in DMSO for about a year proved to be good for arDNA preservation and applicability to PCR. Good DNA preservation for longer periods (several years at room temperature, but preferably at 4°C or -20°C) has also been reported by other researchers (e.g., Dawson *et al.* 1998; Kilpatrick 2002; Hoffman and Amos 2005).

A frequently raised question regards the use of pure-ethanol or IMS for preservation. The study on mackerel gave good preliminary results with IMS, even better than with pure-ethanol for some arDNA extractions and molecular markers (depending also on the percentages of the ethanol solution used for preservation and DNA extraction protocol). However, in my opinion, caution is needed in using IMS preservation because of variability of chemicals that might be present in a particular IMS preparation that might seriously affect extraction of amplifiable DNA. Some researchers have reported difficulties in using IMS-stored biological material for DNA studies (e.g. Boyle *et al.* 2004), but others did not (e.g. Umetsu *et al.* 2002). The use of a chemically pure ethanol solution is recommended, or the use of only IMS for which chemical components are known (e.g. HPLC listing of chemicals present in a particular IMS solution). However, there are reports that even pure-ethanol fixation/preservation resulted in poor yield and degradation of DNA for some organisms (e.g. Reiss *et al.* (1995) for beetles), as well as the existence of some correlation between ethanol-DNA degradation and the extraction (i.e. pre-extraction) procedure applied for isolating DNA from the tissue (Kilpatrick 2002).

Experiments on mackerel tested the relation of ten different preservation methods to the success rate of PCR amplification in conjunction with six DNA extraction protocols and different molecular markers. However, the fidelity of DNA sequences might be affected by different preservation methods. Therefore, sequencing of PCR products is required in order to asses the effect of specimen preservation on sequence fidelity in conjunction with a DNA extraction protocol, genetic marker and genome region.

Collecting and handling of samples is very important for their subsequent usability in molecular work. The time between the death of the organism and its preservation should be very short (i.e. not more than a few minutes) in order to prevent DNA breakage and fragmentation into very small fragments ("the DNA polymers breaking down over a period of hours into fragments only a few hundred nucleotides in length" - Brown 1999). General recommendations for collecting samples are: not keeping dead organism for long before

immersing in fixative (or if immediate fixation is not possible, then keeping dead organisms on ice – cold box, dry ice or liquid nitrogen before fixation), handling specimens with gloves and avoiding direct contact of one individual with another (keeping them in individual containers, or in separate plastic bags in the same container), removing any obvious growth of fungi and plants, or symbiotic and parasitic contaminant organisms on the body of the sample, avoiding the exposure of the collected specimen to sunlight, high temperature or any kind of radiation (i.e. placing collected specimen in a stable environment as soon as possible). Preserved specimens should be stored at constant low temperature (ideally at 4°C, but a low room temperature of about 12-15°C should be sufficient for keeping specimens and their DNA in good conditions). Proper maintenance and monitoring of collection condition (regular pH checking and topping up preservation liquids in particular) is very important for good DNA preservation and for keeping collected samples in good morphological condition. Guidance and recommendations about collecting specimens and maintaining collections for use in molecular studies have already been published by some authors (Prendini *et al.* (2002) is probably the most comprehensive published overview on this subject).

Keeping full records of collected and archived specimens is important for understanding archival DNA (Schander and Halanych 2003). For existing collections, efforts should be made to find the curators' notes of all treatments applied to preserved specimens, including the method of killing animals (for instance, narcotic and relaxing reagents used before fixation, and if used) and chemicals used for maintaining colour and other features of preserved specimens. For newly obtained material, all details since collection (including environmental conditions of surroundings and characteristics of the location) need to be recorded.

For museum specimens used in this study there were no records on (approximate) time between the death of fishes and their fixation, nor records of environmental and other conditions in which fishes were kept before fixation. Neither were there records on topping up preservative during storage and the exact chemical components of these, including pH. The above-mentioned factors might explain the differences seen in amplification of formalin-arDNA extracted from *Nezumia aequalis* and *N. micronychodon* in this study. Formalin-arDNA of *N. aequalis* gave more successful PCRs than formalin-arDNA of *N. micronychodon* despite the fact that the specimens of *N. aequalis* were preserved and stored in the Museum for about 8 years longer than those of *N. micronychodon*. Additionally, these two species came from different waters (*N. micronychodon* from North-West African waters,

whereas *N. aequalis* came from the North-East Atlantic), but there are no records on water composition where these fish specimens came from. Differences in waters, outside temperature and weather conditions in Africa and the North Atlantic are to be expected, but without the exact records, these can only be noted as possible factors. In other words, observed differences in the DNA extractability and PCR usability of *N. aequalis* and *N. micronychodon* are not necessarily caused by different handling of fish specimens when caught, the method of preservation of specimens and the type and quality of chemicals used for fixation and storage. This demonstrates the need for a good understanding of specimen history when working with archival material.

## *4.5. Further directions in Nezumia and macrourid research*

This project had the intention of generally investigating the main factors that might influence recovery of DNA from preserved specimens, i.e. archival material from Steedman's storage solution for which, to my knowledge, there is no available information on optimised DNA extraction protocols or successful DNA investigations. Formalin-fixed, Steedman-preserved *Nezumia* specimens were used as a test material, but there were also attempts to provide molecular data on two *Nezumia* species used in the study (i.e. to make an initial database for further research on these species and the genus *Nezumia*).

A large body of data was accumulated on *Nezumia*: RAPD profiles for selected RAPD primers on the investigated preserved specimens; two sequences from the investigated species with designed *Nezumia*-specific PCR primers and a possible SNP marker. This study provided a list of potentially useful RAPD primers that are applicable for studying *Nezumia* genomes, especially of preserved specimens. The most successful RAPD primers in this study are highly likely to be applicable to other preserved specimens of *Nezumia*.

The sequencing results of this project should make it easier to pursue further genetic research on *Nezumia* species since a starting point is now established for this kind of investigation with the possibility to compare sequences. The developed primers can be used for cross-populations and cross-species experiments in order to confirm the presence and possible variations in targeted sequences of *Nezumia* species and other macrourid fishes, as well as to conduct investigations with other, more distant, fish species in order to investigate whether these primers also can be applied to their genomes.

Further testing on whether the *Nezumia*- specific PCR primers developed (sets: 8, 16 and 18) can be used for generating PCR products from other fish species (closely related and distant) might be informative from the genomics point of view. The initial experiments on mackerel (distantly related to *Nezumia*) have already provided evidence that this sequence was not amplifiable in that species. Investigations towards a better understanding and characterisation of this genome region might be informative and scientifically interesting from many aspects.

The sequence similarity in this 458 bp long (anonymous) region of the genomes of these two species (*Nezumia aequalis* and *N. micronychodon*) might indicate the genetic similarity of these two species in general, but only further molecular investigations (including a "multi-marker approach") could determine the genetic similarity between these two species. Additional molecular markers are required for a more refined analysis of *Nezumia*. New molecular markers might be developed through RAPDs and/or by the use of other molecular techniques and markers. These would be easier to develop on fresh/frozen samples and short-term ethanol and DMSO preserved specimens, but the use of "regular" collections is also possible. Recent developments in deep-sea research and technical developments in collecting deep-sea samples increase the chances of accumulating useful sequences of macrourid fishes, as well as in obtaining fresh/frozen specimens of *Nezumia* species. DNA barcoding of *Nezumia* species should be considered as a priority in the further research.

It is hoped that further studies will benefit from the results presented in this thesis, related to the use of preserved archival collections for molecular investigations and towards creating meaningful guidance for the application of preserved specimens in molecular studies; also that these first molecular data on *Nezumia* will initiate further molecular investigations on this genus and other macrourid fishes.

The main achievements of this project are to have shown that it is feasible to use formalin-fixed, Steedman-preserved archival fish specimens stored for up to 20 years in a museum collection, and to demonstrate that the strategy can be successfully applied in investigations of species lacking any molecular-genetic information. The study also represents a comprehensive overview and development of methodological approaches in order to obtain informative and reliable molecular data from archival collections and preserved specimens. This study also furthers our understanding of DNA degradation in formalin-fixed specimens and it should facilitate further experiments to determine the optimal method for storing

fishes intended for DNA analyses, possible approaches in extracting arDNA from existing collections and their feasibility in molecular studies.

Further investigations on the distribution and type of DNA damage (fragmentation; the presence and distribution of cross-links; point artifactual mutations; insertions/deletions) in mitochondrial and nuclear genomes in differently preserved specimens is important for better understanding and the use of archival DNA. Considering the immense technical advances in molecular biology in the last couple of years, many of these questions might be possible to answer soon, but a multidisciplinary approach is required. As more research is done we will better understand the problems of working with preserved specimens (especially, formalin-preserved ones) and increasingly overcome them.

# Literature cited:


**Abe H., Harada T., Kanehara M., Shimada T., Ohbayashi F. and Oshiki T. (1998a):** Genetic mapping of RAPD markers linked to the densonucleosis refractoriness gene, *nsd-1*, in the silkworm, *Bombyx mori*. *Genes Genet. Syst.* **73**: 237-242.

**Abe H., Kanehara M., Terada T., Ohbayashi F., Shimada T., Kawai S., Suzuki M., Sugasaki T. and Oshiki T. (1998b):** Identification of novel random amplified polymorphic DNAs (RAPDs) on the W chromosome of the domesticated silkworm, *Bombyx mori*, and the wild silkworm, *B. mandarina*, and their retrotransposable element-related nucleotide sequences. *Genes Genet. Syst.* **73**: 243-254.

**Abe H., Ohbayashi F., Shimada T., Sugasaki T., Kawai S., Mita K. and Oshiki T. (2000):** Molecular structure of a novel *gypsy*-Ty3-like retrotransposon (*Kabuki*) and nested retrotransposable elements on the W chromosome of the silkworm *Bombyx mori*. *Mol Gen Genet* **263**: 916-924.

**Abe H., Seki M., Ohbayashi F., Tanaka N., Yamashita J., Fuji T., Yokoyama T., Takahashi M., Banno Y., Sahara K., Yoshido A., Ihara J., Yasukochi Y., Mita K., Ajimura M., Suzuki M.G., Oshiki T. and Shimada T. (2005):** Partial deletions of the W chromosome due to reciprocal translocation in the silkworm *Bombyx mori*. *Insect Molecular Biology* **14**(4): 339-352.

**Abercrombie M., Hickman C.J. and Johnson M.L. (1973):** *A Dictionary of Biology*. Sixth edition. Hunt Bernard Printing Ltd., U.K.

**Aboim M.A., Menezes G.M., Schlitt T. and Rogers A.D. (2005):** Genetic structure and history of populations of the deep-sea fish *Helicolenus dactylopterus* (Delaroche, 1809) inferred from mtDNA sequence analysis. *Molecular Ecology* **14**: 1343-1354.

**Abu Al-Soud W. and Radstrom P. (1998):** Capacity of nine thermostable DNA polymerases to mediate DNA amplification in the presence of PCR-inhibiting samples. *Applied and Enviromental Microbiology* **64**(10): 3748-3753.

**Adams R.P., Zhong M. and Fei Y. (1999):** Preservation of DNA in plant specimens: inactivation and re-activation of DNases in field specimens. *Molecular Ecology* **8**: 681-683.

**Adcock G.J., Ramirez J.H.B., Hauser L., Smith P. and Carvalho G.R. (2000):** Screening of DNA polymorphisms in samples of archived scales from New Zealand snapper. *Journal of Fish Biology* **56**: 1283-1287.

**Agusti N., de Vicente M.C. and Gabarra R. (2000):** Developing SCAR markers to study predation on *Trialeurodes vaporariorum*. *Insect Molecular Biology* **9**(3): 263-268.

**A`Hara S., Harling R., McKinlay R. and Topping C. (1998):** RAPD profiling of spider (Araneae). *The Journal of Arachnology* **26**: 397-400.

**Ahmed M.M.M., Ali B.A. and EI-Zaeem S.W. (2004):** Application of RAPD markers in fish: Part I – some genera (*Tilapia, Sarotherodon and Oreochromis*) and species (*Oreochromis aureus* and *Oreochromis niloticus*) of *Tilapia. International Journal of Biotechnology* **6**(1): 86-93.

**Akbari M., Hansen M.D., Halgunset J., Skorpen F. and Krokan H.E. (2005):** Low copy number DNA template can render polymerase chain reaction error prone in a sequence-dependent manner. *Journal of Molecular Diagnostics* 7(1): 36-39.

**Ali B.A., Huang T-H., Qin D-N. and Wang X-M. (2004):** A review of random amplified polymorphic DNA (RAPD) markers in fish research. *Reviews in Fish Biology and Fisheries* **14**: 443-453.

**Alkami Biosystems, Inc (1999):** ***Alkami Quick Guide for PCR. A laboratory reference for the Polymerase Chain Reaction***. Alkami Biosystems, Inc. I edition, U.S.A.

**Allain V., Biseau A. and Kergoat B. (2003):** Preliminary estimates of French deepwater fishery discards in the Northeast Atlantic Ocean. *Fisheries Research* **60**(1): 185-192.

**Almeida F.S., Sodre L.M.K. and Contel E.P.B. (2003):** Population structure analysis of *Pimelodus maculatus* (Pisces, Siluriformes) from the Tiete and Paranapanema rivers (Brazil). *Genetics and Molecular Biology* **26**(3): 301-305.

**Alonso A., Alves C., Suarez-Mier M.P., Albarran C., Pereira L., Simon L.F., Martin P., Garcia O., Gusmao L., Sancho M. and Amorim A. (2005):** Mitochondrial DNA haplotyping revealed the presence of mixed up benign and neoplastic tissue sections from two individuals on the same prostatic biopsy slide. *J Clin Pathol* **58**: 83-86.

**An S.F. and Fleming K.A. (1991):** Removal of inhibitor(s) of the polymerase chain reaction from formalin fixed, paraffin wax embedded tissues. *J Clin Pathol* 44(11): 924-927.

**Anami Y., Takeuchi T., Mase K., Yasuda J., Hirohashi S., Perucho M. and Noguchi M. (2000):** Amplotyping of microdissected, methanol-fixed lung carcinoma by arbitrarily primed polymerase chain reaction. *Int. J. Cancer (Pred. Oncol.)* **89**: 19-25.

**Andre P., Kim A., Khrapko K. and Thilly W.G. (1997):** Fidelity and mutational spectrum of *Pfu* DNA polymerase on a human mitochondrial DNA sequence. *Genome Research* **7**: 843-852.

**Angel M.V. (2003):** The pelagic environment of the open ocean. pp. 39-79 *in* Tyler P.A. (ed.): ***Ecosystems of the deep oceans*** (*Ecosystems of the World* **28**). First edition. Elsevier Science B.V.

**Antunes A. and Ramos M.J. (2005):** Discovery of a large number of previously unrecognized mitochondrial pseudogenes in fish genomes. *Genomics* **86**: 708-717.

**Aranishi F. (2006):** Single fish egg DNA extraction for PCR amplification. *Conservation Genetics* 7: 153-156.

**Arezi B., Xing W., Sorge J.A. and Hogrefe H.H. (2003):** Amplification efficiency of thermostable DNA polymerases. *Analytical Biochemistry* **321**: 226-235.

**Arroyo-Pardo E., Ocana M.A., Arroyo J.J., Perez-Perez A., Turbon D., Trancho G., Rodriguez-Albarran M.S., Casas J.D., Robledo B. and Bandres F. (2002):** HPLC and UV-spectrophotometry examination of ancient DNA and PCR inhibitors in old human remains. *Ancient Biomolecules* 4(1): 33-41.

**Ashley M.V. and Dow B.D. (1994):** The use of microsatellite analysis in population biology: background, methods and potential applications. Pages 185-201 *in:* Schierwater B., Streit B., Wagner G.P. and DeSalle R., editors. ***Molecular ecology and evolution: Approaches and applications***. Birkhauser Verlag, Basel, Boston, Berlin

**Audic S. and Beraud-Colomb E. (1997):** Ancient DNA is thirteen years old. *Nature Biotechnology* **15**: 855-858.

**Austin J.J., Ross A.J., Smith A.B., Fortey R.A. and Thomas R.H. (1997):** Problems of reproducibility – does geologically ancient DNA survive in amber-preserved insects? *Proc. R. Soc. Lond. B* **264**: 467-474.

**Austin J.J., Arnold E.N. and Bour R. (2003):** Was there a second adaptive radiation of giant tortoises in the Indian Ocean? Using mitochondrial DNA to investigate speciation and biogeography of *Aldabrachelys* (Reptilia, Testudinidae). *Molecular Ecology* **12**: 1415-1424.

**Austin J.J. and Melville A. (2006):** Incorporating historical museum specimens into molecular systematic and conservation genetics research. *Molecular Ecology Notes* **6**(4): 1089-1092.

**Avise J.C. (1978):** RFLP. *The XIV$^{th}$ International Congress of Genetics, Moscow, August 21-30.*

**Avise J.C. (2004):** ***Molecular markers, natural history and evolution***. Second edition. Sinauer Associates, Inc. Publishers Sunderland, Massachusetts, USA.

**Avise J.C., Giblin-Davidson C., Laerm J., Patton J.C. and Lansman R.A. (1979):** Mitochondrial DNA clones and matriarchal phylogeny within and among geographic populations of the pocket gopher, *Geomys pinetis*. *Proc. Natl. Acad. Sci. USA* **76**: 6694-6698.

**Avise J.C., Niegel J.E. and Arnold J. (1984):** Demographic influences on mitochondrial DNA lineage survivorship in animal populations. *J. Mol. Evol.* **20**: 99-105.

**Avise J.C., Helfman G.S., Saunders N.C. and Hales L.S. (1986):** Mitochondrial DNA differentation in North Atlantic eels: population genetic consequences of an unusual life history. *Proc. Natl. Acad. Sci. USA* **83**: 4350-4354.

**Ayyadevara S., Thaden J.J. and Reis R.J.S. (2000):** Anchor polymerase chain reaction disply: A high-trhroughput method to resolve, score, and isolate dimorphic genetic markers based on interspersed repetitive DNA elements. *Analytical Biochemistry* **284**: 19-28.

**Azeredo-Espin A.M.L. and Lessinger A.C. (2006):** Genetic approaches for studying myiasis-causing flies: molecular markers and mitochondrial genomics. *Genetica* **126**: 111-131.

**Bailey C.D., Hughes C.E. and Harris S.A. (2004):** Using RAPDs to identify DNA sequence loci for species level phylogeny reconstruction: an example from *Leucaena* (Fabaceae). *Systematic Botany* **29**(1): 4-14.

**Bailey D.M., Jamieson A.J., Bagley P.M., Collins M.A. and Priede I.G. (2002):** Measurement of in situ oxygen consumption of deep-sea fish using an autonomous lander vehicle. *Deep-Sea Research I* **49**: 1519-1529.

**Baker R.J. (1994):** Some thoughts on conservation, biodiversity, museums, molecular characters, systematics, and basic research. *Journal of Mammalogy* **75**(2): 277-287.

**Balaban R.S., Nemoto S. and Finkel T. (2005):** Mitochondria, oxidants, and aging. *Cell* **120**: 483-495.

**Ball S.L., Hebert P.D.N., Burian S.K. and Webb J.M. (2005):** Biological identifications of mayflies (Ephemeroptera) using DNA barcodes. *J. N. Am. Benthol. Soc.* **24**(3): 508-524.

**Banerjee M. and Brown T.A. (2002):** Differential preservation of nuclear and chloroplast DNA? *Ancient Biomolecules* **4**(3): 122. (*Abstracts of the 6th International Conference of Ancient DNA and Associated Biomolecules, Tel Aviv, Israel, July 21-25, 2002).*

**Banerjee M. and Brown T.A. (2004):** Non-random DNA damage resulting from heat treatment: implications for sequence analysis of ancient DNA. *Journal of Archaeological Science* **31**(1): 59-63.

**Banki K., Halladay D. and Perl A. (1994):** Cloning and expression of the human gene for transaldolase. *The Journal of Biological Chemistry* **269**(4): 2847-2851.

**Banks M.A., Waters C. and Hedgecock D. (1993):** Discrimination between closely related Pacific oyster species (*Crassostrea*) via mitochondrial DNA sequences coding for large subunit rRNA. *Molecular Marine Biology and Biotechnology* **2**: 129-136.

**Baquero F. (2005):** Evolution and the nature of time. *International Microbiology* **8**: 81-91.

**Baratti M., Bazzicalupo M., de Filippo C. and Messana G. (1999):** Detection of genetic variability in stygobitic isopods using RAPD markers. *Crustaceana* **72**(7): 625-634.

**Baratti M., Argano R., Burchi A. and Messana G. (2003):** Analysis of genetic variability of stygobitic isopod populations (*Stenasellus racovitzai,* Isopoda) with RAPD markers and comparison to allozyme data. *Crustaceana* **76**(1): 39-47.

**Bardakci F. (2001):** Random Amplified Polymorphic DNA (RAPD) markers. *Turk J Biol* **25**: 185-196.

**Barker S., Murray D., Zheng J., Li L. and Weinfeld M. (2005a):** A method for the isolation of covalent DNA-protein crosslinks suitable for proteomics analysis. *Analytical Biochemistry* **344**(2): 204-215.

**Barker S., Weinfeld M. and Murray D. (2005b):** DNA-protein crosslinks: their induction, repair, and biological consequences. *Mutation Research* **589**: 111-135.

**Barret P., Delourme R., Foisset N. and Renard M. (1998):** Development of a SCAR (sequence characterised amplified region) marker for molecular tagging of the dwarf BREIZH (*Bzh*) gene in *Brassica napus* L. *Theoretical and Applied Genetics* **97**(5-6): 828-833.

**Barrett R.D.H. and Hebert P.D.N. (2005):** Identifying spiders through DNA barcodes. *Can. J. Zool.* **83**: 481-491.

**Baum D.A. (1998):** Individuality and the existence of species through time. *Syst. Biol.* **47**(4): 641-653.

**Beattle A.J. and Ollver I. (1994):** Taxonomic minimalism. *TREE* **9**(12): 488-490.

**Beck J.L. and Hyman B.C. (1988):** Role of sequence amplification in the generation of nematode mitochondrial DNA polymorphism. *Current Genetics* **14**(6): 627-636.

**Beebee T.J.C., Rowe G. and Burke T. (1998):** Archive contributions to a molecular phylogeography of the toad *Bufo calamita* in Britain. *Biochemical Genetics* **36**: 219-228.

**Beebee T.J.C., Rowe G. and Arntzen W. (1999):** Identification of newts (*Triturus* sp.) and their hybrids using molecular methods. *J. Zool. Lond.* **249**: 43-47.

**Behura S.K., Sahu S.C., Rajamani S., Devi A., Mago R., Nair S. and Mohan M. (1999):** Differentiation of Asian rice gall midge, *Orseolia oryzae* (Wood-Mason), biotypes by sequence characterized amplified regions (SCARs). *Insect Molecular Biology* **8**(3): 391-397.

**Behura S.K., Valicente F.H., Rider S.D. Jr., Shun-Chen M., Jackson S. and Stuart J.J. (2004):** A physically anchored genetic map and linkage to avirulence reveals recombination suppression over the proximal region of Hessian fly chromosome A2. *Genetics* **167**: 343-355.

**Benton M.J. (2000):** Stems, nodes, crown clades, and rank-free lists: is Linnaeus dead? *Biological Reviews* **75**: 633-648.

**Benzie J.A.H. (1999):** Major genetic differences between crown-of-thorns starfish (Acanthaster-Planci) populations in the Indian and Pacific Oceans. *Evolution* **53**(6): 1782-1795.

**Berg W.J. and Ferris S.D. (1984):** Restriction endonuclease analysis of salmonid mitochondrial DNA. *Can. J. Fish. Aquat. Sci.* **41**: 1041-1047.

**Berger A., Bruschek M., Grethen C., Sperl W. and Kofler B. (2001):** Poor storage and handling of tissue mimics mitochondrial DNA depletion. *Diagnostic Molecular Pathology* **10**(1): 55-59.

**Bernstein J.L., Thompson W. D., Casey G., DiCioccio R.A., Whittemore A.S., Diep A.T., Thakore S.S., Vaziri S., Xue S. and Haile R.W. (2002):** Comparison of techniques for the successful detection of *BRCA 1* mutations in fixed paraffin-embedded tissue. *Cancer Epidemiology, Biomarkers and Prevention* **11**: 809-814.

**Bertozzini E., Penna A., Pierboni E., Bruce I. and Magnani M. (2005):** Development of new procedures for the isolation of phytoplankton DNA from fixed samples. *Journal of Applied Phycology* **17**: 223-229.

**Beznosov V.N. and Suzdaleva A.L. (2001):** The effect of upwelling on zooplankton and fish larvae in the sea surface layer. *Oceanology* **41**(6): 833-836.

**Bhadury P., Austen M.C., Bilton D.T., Lambshead J.D., Rogers A.D. and Smerdon G.R. (2005):** Combined morphological and molecular analysis of individual nematodes through short-term preservation in formalin. *Molecular Ecology Notes* **5**: 965-968.

**Bhadury P., Austen M.C., Bilton D.T., Lambshead J.D., Rogers A.D. and Smerdon G.R. (2006a):** Exploitation of archived marine nematodes – a hot lysis DNA extraction protocol for molecular studies. *Zoologica Scripta* (Online Early Artical).

**Bhadury P., Austen M.C., Bilton D.T., Lambshead J.D., Rogers A.D. and Smerdon G.R. (2006b):** Development and evaluation of a DNA-barcoding approach for the rapid identification of nematodes. *Marine Ecology Progress Series* **320**: 1-9.

**Bielawski J.P. and Pumo D.E. (1997):** Randomly amplified polymorphic DNA (RAPD) analysis of Atlantic Coast striped bass. *Heredity* **78**: 32-40.

**Bilodeau A.L., Lankford W.S., Kim T.J., Felder D.L. and Neigel J.E. (1999):** An ultrasensitive method for detection of single crab larvae (*Sesarma reticulatum*) by PCR amplification of a highly repetetive DNA sequence. *Molecular Ecology* **8**: 683-684.

**Binladen J., Wiuf C., Gilbert M.T.P., Bunce M., Barnett R., Larson G., Greenwood A.D., Haile J., Ho S.Y.W., Hansen A.J. and Willerslev E. (2006):** Assessing the fidelity of ancient DNA sequences amplified from nuclear genes. *Genetics* **172**: 733-741.

**Birky C.W.J. (1995):** Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proc. Natl. Acad. Sci. USA* **92**: 11331-11338.

**Birky C.W. Jr, Fuerst P. and Maruyama T. (1989):** Organelle gene diversity under migration, mutation, and drift: equilibrium expectations approach to equilibrium, effects of heteroplasmic cells and comparison to nuclear genes. *Genetics* **121**: 613-627.

**Birmeta G., Nybom H. and Bekele E. (2004):** Distinction between wild and cultivated enset (*Ensete ventricosum*) gene pools in Ethiopia using RAPD markers. *Hereditas* **140**: 139-148.

**Biscoito M., Segonzac M. and Almeida A.J. (2001):** New zoarcid fish species from deep-sea hydrothermal vents of the Atlantic and Pacific Oceans. *InterRidge News* **10**(1): 15-17.

**Blaber M. (1998):** BCH5425 Molecular Biology and Biotechnology. (http://wine1.sb.fsu.edu/bch5425/lect13/lect13.htm)

**Black A.R. and Dodson S.I. (2003):** Ethanol: a better preservation technique for *Daphnia*. *Limnology and Oceanography: Methods* **1**: 45-50.

**Blaxter M. L. (2004):** The promise of a molecular taxonomy. *Phil. Trans. R. Soc. B* **359**: 669–679.

**Bogenhagen D.F., Rousseau D. and Burke S. (2008):** The layered structure of human mitochondrial DNA nucleoids. *J. Biol. Chem.* **283**(6): 3665-3675.

**Boman J., Gaydos C.A. and Quinn T.C. (1999):** Molecular diagnosis of *Chlamydia pneumoniae* infection. *Journal of Clinical Microbiology* **37**(12): 3791-3799.

**Bond J.E. (2004):** Systematics of the Californian euctenizine spider genus *Apomastus* (Araneae: Mygalomorphae: Cyrtaucheniidae): the relationship between molecular and morphological taxonomy. *Invertebrate Syst.* **18**: 361-376.

**Bonin S., Petrera F., Niccolini B. and Stanta G. (2003):** PCR analysis in archival postmortem tissues. *J Clin Pathol: Mol Pathol* **56**: 184-186.

**Boom R., Sol C.J.A., Salimans M.M.M., Jansen C.L., Wertheim-van Dillen P.M.E. and van der Noordaa J. (1990):** Rapid and simple method for purification of nucleic acids. *J. Clinical Microbiol.* **28**: 495-503.

**Boore J.L. (1999):** Animal mitochondrial genomes. *Nucleic Acids Research* **27**(8): 1767-1780.

**Boore J.L., Macey J. and Medina M. (2005):** Sequencing and comparing whole mitochondrial genomes of animals. *Part B: Producing the biochemical data II. Methods in Enzymology* **395**: 311-348

**Bottger-Schnack R., Hagen W. and Schnack-Schiel S.B. (2001):** The microcopepod fauna in the Gulf of Aqaba, northern Red Sea: species diversity and distribution of Oncaeidae (Poecilostomatoida). *Journal of Plankton Research* **23**(9): 1029-1035.

**Bouwman A.S. and Brown T.A. (2002):** Comparison between silica-based methods for the extraction of DNA from human bones from 18$^{th}$ to mid-19$^{th}$ Century London. *Ancient Biomolecules* 4(4): 173-178.

**Bouwman A.S. and Brown T.A. (2005):** The limits of biomolecular palaeopathology: ancient DNA cannot be used to study venereal syphilis. *Journal of Archaeological Science* **32**: 703-713.

**Bouzat J.L., Lewin H.A. and Paige K.N. (1998):** The ghost of genetic diversity past: historical DNA analysis of the greater prairie chicken. *The American Naturalist* **152**(1): 1-6.

**Boyce T.M., Zwick M.E. and Aquadro C.F. (1989):** Mitochondrial DNA in the bark weevils: size, structure and heteroplasmy. *Genetics* **123**: 825-836.

**Boyle E.E., Zardus J.D., Chase M.R., Etter R.J. and Rex M.A. (2004):** Strategies for molecular genetic studies of preserved deep-sea macrofauna. *Deep-Sea Research I* **51**: 1319-1336.

**Branco M., Ferrand N. and Monnerot M. (2000):** Phylogeography of the European rabbit (*Oryctolagus cuniculus*) in the Iberian Peninsula inferred from RFLP analysis of the cytochrome *b* gene. *Heredity* **85**(4): 307-317.

**Brandstatter A., Sanger T., Lutz-Bonengel S., Parson W., Beraud-Colomb E., Wen B., Kong Q-P., Bravi C.M., Bandelt H-J. (2005):** Phantom mutation hotspots in human mitochondrial DNA. *Electrophoresis* **26**: 3414-3429.

**Brandt A. (2005):** Evolution of Antarctic biodiversity in the context of the past: the importance of the Southern Ocean deep sea. *Antarctic Science* **17**: 509-521.

**Braslavsky I., Hebert B., Kartalov E. and Quake S.R. (2003):** Sequence information can be obtained from single DNA molecules. *PNAS* **100**(7): 3960-3964.

**Bredel M., Bredel C., Juric D., Kim Y., Vogel H., Harsh G.R., Recht L.D., Pollack J.R. and Sikic B.I. (2005):** Amplification of whole tumor genomes and gene-by-gene mapping of genomic aberrations from limited sources of fresh-frozen and paraffin-embedded DNA. *Journal of Molecular Diagnostics* **7**(2): 171-182.

**Brenke N., Brix S. and Knuschke T. (2005):** A new deep-sea isopod species from the Angola Basin: *Prochelator angolensis* sp. nov. (Asellota: Desmosomatidae). *Organisms Diversity and Evolution* **5**(Suppl. 1): 179-188.

**Briggs D.E.G. (1999):** Molecular taphonomy of animal and plant cuticles: selective preservation and diagenesis. *Phil. Trans. R. Soc. Lond. B* **354**: 7-17.

**Britten R.J. and Kohne D. (1968):** Repeated sequences in DNA. *Science* **161**: 529-540.

**Brohede J. and Ellegren H. (1999):** Microsatellite evolution: polarity of substitutions within repeats. *Proc. R. Soc. Lond. B* **266**: 825-833.

**Brower A.V.Z. (1995):** Taxonomic minimalism. *TREE* **10**(5): 203.

**Brown J.W., Miller S.E. and Horak M. (2003):** Studies on New Guinea moths. 2. Description of a new species of *Xenothictis* Meyrick (Lepidoptera: Tortricidae: Archipini). *Proc. Entomol. Soc. Wash.* **105**(4): 1043-1050.

**Brown K.A. (2001):** The future is bright – directions for future research with ancient DNA in forensic archaeology. *Ancient Biomolecules* **3**(3): 159-166.

**Brown T.A. (1999):** Genetic material. pp. 133-138 *in* Carter D. and Walker A.K., edtors ***Care and conservation of natural history collections***. Butterworth-Heinemann, Oxford (The Reed Elviser plc group; Reed Educational and Profesional Publishing Ltd 1999).

**Brown T.A. (2001):** ***Gene cloning and DNA analysis: An introduction***. Fourth edition. Blackwell Science Ltd.

**Brown T.A. (2002):** ***Genomes II***. BIOS, Oxford.

**Brown T.A., Allaby R.G., Brown K.A., O'Donoghue and Sallares R. (1994):** DNA in wheat seeds from European archaeological sites. *Experientia* **50**: 571-575.

**Brown T.A., Allaby R.G., Sallares R. and Glynis J. (1998):** Ancient DNA in charred wheats: Taxonomic identification of mixed and single grains. *Ancient Biomolecules* **2**(2/3): 185-194.

**Brown W.M, George M. jun. and Wilson A.C. (1979):** Rapid evolution of mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* **76**: 1967-1971.

**Brown W.M. (1985):** The mitochondrial genome of animals. *In:* R.J. MacIntyre, editor. ***Molecular evolutionary genetics.*** Plenum Press, New York.

**Brzuzan P., Ciesielski S., Kusznierz J. and Jurczyk L. (2004):** Identification of whitefish remains by mtDNA analysis. *Ann. Zool. Fennici* **41**: 61-68.

**Buchan J.C., Archie E.A., van Horn R.C., Moss C.J. and Alberts S.C. (2005):** Locus effects and sources of error in noninvasive genotyping. *Molecular Ecology Notes* **5**: 680-683.

**Buckleton J., Triggs C.M. and Walsh S.J. (2005):** ***Forensic DNA evidence interpretation***. CRC Press, Boca Raton, London, New York, Washington.

**Bucklin A. and Allen L.D. (2004):** MtDNA sequencing from zooplankton after long-term preservation in buffered formalin. *Molecular Phylogenetics and Evolution* **30**(3): 879-882.

**Budowle B., Allard M.W., Wilson M.R. and Chakraborty R. (2003):** Forensic and mitochondrial DNA: Applications, debates, and foundations. *Annu. Rev. Genomics Hum. Genet.* **4**: 119-141.

**Bull J.J., Huelsenbeck J.P., Cunningham C.W., Swofford D.L. and Waddell P.J. (1993):** Partitioning and combining data in phylogenetic analysis. *Syst. Biol.* **42**: 384-397.

**Burger G., Forget L., Zhu Y., Gray M.W. and Lang B.F. (2003a):** Unique mitochondrial genome architecture in unicellular relatives of animals. *PNAS* **100**(3): 892-897.

**Burger G., Gray M.W. and Lang B.F. (2003b):** Mitochondrial genomes: anything goes. *TRENDS in Genetics* **19**(12): 709-716.

**Buroker N.E., Brown J.R., Gilbert T.E., O`Hara P.J., Beckenbach A.T., Thomas W.K. and Smith M.J. (1990):** Length heteroplasmy of sturgeon mitochondrial DNA: an illegitimate elongation model. *Genetics* **124**: 157-163.

**Busby M.S. (2005):** An unusual macrourid larva (Gadiformes) from San Juan Island, Washington, USA. *Ichthyological Research* **52**: 86-89.

**Cadet J., Berger M., Douki T. and Ravanat J.-L. (1997):** Oxidative damage to DNA: formation, measurement, and biological significance. *Reviews of Physiology, Biochemistry and Pharmacology* **131**: 1-87.

**Caetano-Anolles G., Bassam B.J. and Gresshoff P.M. (1992):** Primer-template interactions during DNA amplification fingerprinting with single arbitrary oligonucleotides. *Molecular and General Genetics* **235**: 157-165.

**Campbell I.C. (1995):** Taxonomic minimalism. *TREE* **10**(5): 203.

**Cano M.J., Werner O. and Guerra J. (2005):** A morphometric and molecular study in *Tortula subulata* complex (Pottiaceae, Bryophyta). *Botanical Journal of the Linnean Society* **149**: 333-350.

**Cano R.J. and Poinar H.N. (1993):** Rapid isolation of DNA from fossil and museum specimens sutiable for PCR. *BioTechniques* **15**: 432-435.

**Cantatore P. and Saccone C. (1987):** Organization, structure, and evolution of mammalian mitochondrial genes. *Int. Rev. Cytol.* **180**: 149-208.

**Cantino P. and de Queiroz K. (2004):** Phylocode: A phylogenetic code of biological nomenclature. http://www.ohiou.edu/phylosode/

**Cao D. and Card J.H. (1997):** Pedigree and RAPD-based DNA analysis of commercial U.S. rice cultivars. *Crop Science* **37**(5): 1630-1635.

**Carter J. (1997):** An investigation of pH changes in a selection of formaldehyde buffering agents used on a fish, parasitology research collection. *Collection Forum* **13**(1): 1-10.

**Carter J.D. (2003):** *The effects of preservation and conservation treatments on the DNA of museum invertebrate fluid preserved collections.* M.Phil Thesis, University of Wales, Cardiff, UK.

**Carvalho A.O.R. and Vieira L.G.E. (2000):** Comparison of preservation methods of *Atta* spp. (Hymenoptera: Formicidae) for RAPD analysis. *An. Soc. Entomol. Brasil* **29**(3): 489-496.

**Cattaneo C., Craig O.E. and James N.T. (1997):** Comparison of three DNA extraction methods on bone and blood stains up to 43 years old and amplification of three different gene sequences. *J. Forensic Sci* **42** (6): 1126-1235.

**Cawkwell L. and Quirke P. (2000):** Direct multiplex amplification of DNA from a formalin fixed, paraffin wax embedded tissue section. *J Clin Pathol: Mol Pathol* **53**: 51-52.

**Chakraborty A., Sakai M. and Iwatsuki Y. (2006):** Museum fish specimens and molecular taxonomy: A comparative study on DNA extraction protocols and preservation techniques. *J. Appl. Ichthyol.* **22**: 160-166.

**Chang Y.-T. and Loew G.H. (1994):** Reaction mechanisms of formaldehyde with endocyclic imino groups of nucleic acid bases. *J. Am. Chem. Soc.* **116**: 3548-3555.

**Chase M.R., Etter R.J., Rex M.A. and Quattro J.M. (1998a):** Extraction and amplification of mitochondrial DNA from formalin-fixed deep-sea molluscs. *BioTechniques* **24**(2): 243-247.

**Chase M.R., Etter R.J., Rex M.A. and Quattro J.M. (1998b):** Bathymetric patterns of genetic variation in a deep-sea protobranch bivalve, *Deminucula atacellana. Marine Biology* **131**: 301-308.

**Chaw Y.F.M., Crane L.E., Lange P. and Shapiro R. (1980):** Isolation and identification of cross-links from formaldehyde-treated nucleic acids. *Biochemistry* **19**: 5525-5531.

**Chelomina G.N. (2006):** Ancient DNA. *Russian Journal of Genetics* **42**(3): 219-233.

**Chiari L. and Sodre L.M.K. (2001):** Study of eight species of the Anostomidae family (Pisces, Characiformes) by RAPD analysis. *Acta Scientiarum: Maringa* **23**(2): 445-451.

**Chiou M-L., Shao K-T. and Iwamoto T. (2004a):** A new species, *Caelorinchus sheni*, and 19 new records of grenadiers (Pisces: Gadiformes: Macrouridae). *Zoological Studies* **43**(1): 35-50.

**Chiou M-L., Shao K-T. and Iwamoto T. (2004b):** New species of *Caelorinchus* (Macrouridae, Gadiformes, Teleostei) from Taiwan, with a redescription of *Caelorinchus brevirostris* Okamura. *Copeia* **2004**(2): 298-304.

**Chu K.H., Li C.P. and Qi J. (2006):** Ribosomal RNA as molecular barcodes: a simple correlation analysis without sequence alignment. *Bioinformatics* **22**(14): 1690-1701.

**Ciesielski S., Brzuzan P. and Luczynski M. (2002):** Recovery and analysis of mitochondrial DNA from ancient bones of common bream, *Abramis brama* L. *Ancient Biomolecules* **4**(1): 43-46.

**Cipollaro M., Galderisi U. and Di Bernardo G. (2005):** Ancient DNA as a multidisciplinary experience. *Journal of Cellular Physiology* **202**: 315-322.

**Clark A.G., and Lanigan M.S. (1993):** Prospects for estimating nucleotide divergence with RAPD`s. *Molecular Biology and Evolution* **10**: 1096-1111.

**Clark J.M. (1988):** Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. *Nucleic Acids Res.* **16**: 9677–9686.

**Clark M.S., Clarke A., Cockell C.S., Convey P., Detrich III H.W., Fraser K.P.P., Johnston I.A., Methe B.A., Murray A.E., Peck L.S., Romisch K. and Rogers A.D. (2004):** Antarctic genomics. *Comparative and Functional Genomics* **5**: 230-238.

**Clay O., Douady C.J., Carels N., Hughes S., Bucciarelli G. and Bernardi G. (2003):** Using analytical ultracentrifugation to study compositional variation in vertebrate genomes. *Eur. Biophys. J.* **32**: 418-426.

**Coggan R.A., Gordon J.D.M. and Merrett N.R. (1999):** Aspects of the biology of *Nezumia aequalis* from the continental slope west of the British Isles. *Journal of Fish Biology* **54**: 152-170.

**Cohen D.M., Inada T., Iwamoto T. and Scialabba N. (1990):** *FAO species catalogue, Vol. 10. Gadiformes fishes of the world (Order: Gadiformes).* An annotated and illustrated catalogue of codes, hakes, grenadiers and other gadiformes fishes known to date. *FAO Fisheries Synopsis* No. 125.

**Cohen S.N., Chang A.C.Y. and Hsu L. (1972):** Nonchromosomal antibiotic resistance in bacteria: Genetic transformation of *Escherichia coli* by R-Factor DNA. *Proc. Nat. Acad. Sci. USA* **69**(8): 2110-2114.

**Colborn J., Crabtree R.E., Shaklee J.B., Pfeiler E. and Bowen B.W. (2001):** The evolutionary enigma of bonefishes (*Albula* spp.): Cryptic species and ancient separations in a globally distributed shorefish. *Evolution* **55**(4): 807-820.

**Collette B.B. and Vecchione M. (1995):** Interactions between fisheries and systematics. *Fisheries* (Systematics) **20**(1): 20-25.

**Collins D., Mill R.R. and Moller M. (2003):** Species separation of *Taxus baccata*, *T. canadensis* and *T. cuspidata* (Taxaceae) and origins of their reputed hybrids inferred from RAPD and cpDNA data. *American Journal of Botany* **90**(2): 175-182.

**Collins M.A., Bailey D.M., Ruxton G.D. and Priede I.G. (2005):** Trends in body size across an environmental gradient: A differential response in scavenging and non-scavenging demersal deep-sea fish. *Proceedings of the Royal Society – Biological Sciences (Series B)* **272**(1576): 2051-2057.

**Collins T.M., Trexler J.C., Nico L.G. and Rawlings T.A. (2002):** Genetic diversity in a morphologically conservative invasive taxon: Multiple introductions of swamp eels to the Southeastern United States. *Conservation Biology* **16**(4): 1024-1035.

**Comeron J.M. (2001):** What controls the length of noncoding DNA? *Current Opinion in Genetics and Development* **11**: 652-659.

**Congiu L., Chicca M., Cella R., Rossi R. and Bernacchia G. (2000):** The use of random amplified polymorphic DNA (RAPD) markers to identify strawberry varieties: a forensic application. *Molecular Ecology* **9**: 229-232.

**Cook L.M., Soltis P.S., Brunsfeld S.J. and Soltis D.E. (1998):** Multiple independent formations of *Tragopogon* tetraploids (Asteraceae): evidence from RAPD markers. *Molecular Ecology* **7**: 1293-1302.

**Coombs N.J., Gough A.C. and Primrose J.N. (1999):** Optimisation of DNA and RNA extraction from archival formalin-fixed tissue. *Nucleic Acids Research* **27**(16): e12.

**Cooper A. (1994):** Dried samples: soft tissues, DNA from museum specimens. pp. 149-165. *In:* Herrmann B. and Hummed S. (eds): ***Ancient DNA.*** Springer, New York.

**Cooper A. (2006):** The Year of Mammoth. *PLoS Biology* **4**(3): e78.

**Cooper A., Poinar H.N., Paabo S., Radovcic J., Debenath A., Caparros M., Barroso-Ruiz C., Bertranpetit J., Nielsen-Marsh C., Hedges R.E.M. and Sykes B. (1997):** Neanderthal genetics. *Science* **277**: 1021-1024.

**Cooper A. and Poinar H.N. (2000):** Ancient DNA: Do it right or not at all. *Science* **289**: 1139.

**Cooper A., Lalueza-Fox C., Anderson S., Rambaut A., Austin J. and Ward R. (2001):** Complete mitochondrial genome sequences of two extinct moas clarify ratite evolution. *Nature* **409**: 704-707.

**Corthals A. and DeSalle R. (2005):** An application of tissue and DNA banking for genomics and conservation: The Ambrose Monell cryo-collection (AMCC). *Syst. Biol.* **54**(5): 819-823.

**Cossins A.R. and Macdonald A.G. (1989):** The adaptation of biological membranes to temperature and pressure: fish from the deep and cold. *Journal of Bioenergetics and Biomembranes* **21**: 115-135.

**Costa F.O, Cunha M.R., Neuparth T., Theodorakis C.W., Costa M.H. and Shugart L.R. (2004):** Application of RAPD DNA fingerprinting in taxonomic identification of amphipods: a case-study with *Gammarus* species (Crustacea: Amphipoda). *Journal of the Marine Biological Association of the UK* **84**: 171-178.

**Coura R., Prolla J.C., Meurer L. and Ashton-Prolla P. (2005):** An alternative protocol for DNA extraction from formalin fixed and paraffin wax embedded tissue. *J Clin Pathol* **58**: 894-895.

**Cowen R.K., Lwiza K.M.M., Sponaugle S., Paris C.B. and Olson D.B. (2000):** Connectivity of marine populations: Open or closed? *Science* **287**: 857-859.

**Cozzolino S., Cafasso D., Pellegrino G., Musacchio A. and Widmer A. (2006):** Genetic variation in time and space: the use of herbarium specimens to reconstruct patterns of genetic variation in the endangered orchid *Anacamptis palustris*. *Conservation Genetics* (Online Early Artical).

**Creasey S., Rogers A.D., Tyler P., Young C. and Gage J. (1997):** The population biology and genetics of the deep-sea spider crab, *Encephaloides armstrongi* Wood-Mason 1891 (Decapoda: Majidae). *Phil. Trans. R. Soc. Lond. B* **352**: 365-379.

**Creasey S.S. and Rogers A.D. (1999):** Population genetics of bathyal and abyssal organisms. *Advances in Marine Biology* **35**: 1-151.

**Crimmen O.A. (1989):** Phenoxetol: An unsatisfactory preservative for fishes. *Newsletter, Biology Curators Group* **5**(3): 26-27

**Crisan D. and Mattson J.C. (1993):** Retrospective DNA analysis using fixed tissue specimens. *DNA Cell Biol* **12**: 455-464.

**Criscuolo G. (1994):** Museum spirit collection and the preservation of DNA. *Conservation News,* No. 54: 39-40.

**Cronin T.M. and Raymo M.E. (1997):** Orbital forcing of deep-sea benthic species diversity. *Nature* **385**(6617): 624-627.

**Cummins J.M., Wakayama T. and Yanagimachi R. (1998):** Fate of microinjected spermatid mitochondria in the mouse oocyte and embryo. *Zygote* **6**: 213-222.

**Curole J.P. and Kocher T.D. (1999):** Mitogenomics: digging deeper with complete mitochondrial genomes. *TREE* **14**(10): 394-398.

**Dai H., Lo Y-S., Litvinchuk A., Wang Y-T., Jane W-N., Hsiao L-J. and Chiang K-S. (2005):** Structural and functional characterizations of mung bean mitochondrial nucleoids. *Nucleic Acids Research* **33**(15): 4725-4739.

**Das M., Bhattacharya S. and Pal A. (2005):** Generation and characterization of SCARs by cloning and sequencing of RAPD products: A strategy for species-specific marker development in bamboo. *Annals of Botany* **95**: 835-841.

**Dasmahapatra K.K. and Mallet J. (2006):** DNA barcodes: recent successes and future prospects. *Heredity* **97**: 254-255.

**D`Aurelio M., Gajewski C.D., Lin M.T., Mauck W.M., Shao L.Z., Lenaz G., Moraes C.T. and Manfredi G. (2004):** Heterologous mitochondrial DNA recombination in human cells. *Human Molecular Genetics* **13**(24): 3171-3179.

**Dawson M.N., Raskoff K.A. and Jacobs D.K. (1998):** Field preservation of marine invertebrate tissue for DNA analyses. *Molecular Marine Biology and Biotechnology* 7(2): 145-152.

**Dayton P.K. (2003):** The importance of the natural sciences to conservation. (An American Society of Naturalists Symposium Paper). *The American Naturalist* **162**(1): 1-13.

**Dean M.D. and Ballard W.O. (2001):** Factors affecting mitochondrial DNA quality from museum preserved *Drosophila simulans*. *Entomologia Experimentalis et Applicata* **98**: 279-283.

**De Giorgi C., Sialer M.F. and Lamberti F. (1994):** Formalin-induced infidelity in PCR-amplified DNA fragments. *Molecular and Cell Probes* **8**(6): 459-462.

**de Grey A.D.N.J. (2005):** Forces maintaining organellar genomes: is any as strong as genetic code disparity or hydrophobicit? *BioEssays* **27**: 436-446.

**de Forges B.R., Koslow J.A. and Poore G.C.B. (2000):** Diversity and endemism of the benthic seamount fauna in the southwest Pacific. *Nature* **405**: 944-947.

**de Kloet S.R. (2001):** Development of a CAPS (cleaved amplified polymorphics sequence) assay for sex identification of the emu (*Dromaius novaehollandiae*). *Molecular Ecology Notes* **1**: 273-275.

**DeLaat D.M., Carvalho M.R.S., Lovato M.B., Acedo M.D.P. and da Fonseca C.G. (2005):** Applicability of RAPD markers on silver-stained polyacrylamide gels to ascertain genetic diversity in *Peripatus acacioi* (Peripatidae; Onychophora). *Genetics and Molecular Research* **4**(4): 716-725.

**De Ley P., De Ley I.T., Morris K., Abebe E., Mundo-Ocampo M., Yoder M., Heras J., Waumann D., Rocha-Olivares A., Burr A.H.J., Baldwin J.G. and Thomas W.K. (2005):** An integrated approach to fast and informative morphological vouchering of nematodes for applications in molecular barcoding. *Phil. Trans. R. Soc. B* **360**(1462): 1945-1958.

**Delfour C., Roger P., Bret C., Berthe M-L., Rochaix P., Kalfa N., Raynaud P., Bibeau F., Maudelonde T. and Boulle N. (2006):** RCL2, a new fixative, preserves morphology and nucleic acid integrity in paraffin-embedded breast carcinoma and microdissected breast tumor cells. *Journal of Molecular Diagnostics* **8**(2): 157-169.

**Desagher S. and Martinou J-C. (2000):** Mitochondria as the central control point of apoptosis. *Trends in Cell Biology* **10**: 369-377.

**DeSalle R. (1994):** Implications of ancient DNA for phylogenetic studies. *Experientia* **50**(6): 543-550.

**DeSalle R. and Templeton A.R. (1988):** Founder effects and the rate of mitochondrial DNA evolution in Hawaiian *Drosophila*. *Evolution* **42**: 1076-1084.

**DeSalle R., Barcia M. and Wray C. (1993a):** PCR jumping in clones of 30-million-year-old DNA fragments from amber preserved termites (*Mastotermes electrodominicus*). *Experientia* **49**: 906-909.

**DeSalle R., Williams A.K. and George M. (1993b):** Isolation and characterization of animal mitochondrial DNA. *Methods in Enzymology* **224**: 176-204.

**DeSalle R., Egan M.G. and Siddall M. (2005):** The unholy trinity: taxonomy, species delimitation and DNA barcoding. *Phil. Trans. R. Soc. B* **360**: 1905-1916.

**Dessauer H.C. and Hafner M.S., eds. (1984):** ***Collections of Frozen Tissues: Value, Management, Field and Laboratory Procedures, and Directory of Existing Collections***. Assoc. Systematics Collections, Washington, D.C.

**Dessauer H.C., Cole C.J. and Hafner M.S. (1996):** Collection and storage of tissues. pp. 29-47 *in* Hillis D.M., Moritz C. and Mable B.K., editors. ***Molecular systematics***. Second edition. Sinauer Associates Inc., Publishers Sunderland, Massachusetts, USA.

**Deveney M.R. and Whittington I.D. (2001):** A technique for preserving pigmentation in some capsalid monogeneans for taxonomic purposes. *Systematic Parasitology* **48**(1): 31-35.

**DeVries S., Nyante S., Korkola J., Segraves R., Nakao K., Moore D., Bac H., Wilhelm M., Hwang S. and Waldman F. (2005):** Array-based comparative genomic hybridization from formalin-fixed, paraffin-embedded breast tumors. *Journal of Molecular Diagnostics* 7(1): 65-71.

**De Wolf H., Blust R. and Backeljau T. (2004):** The use of RAPD ecotoxicology. *Mutation Research* **566**: 249-262.

**Dewoody J., Nason J.D. and Hipkins V.D. (2006):** Mitigating scoring errors in microsatellite data from wild populations. *Molecular Ecology Notes* **6**(4): 951-957.

**Diaz-Viloria N., Sanchez-Velasco L. and Perez-Enriquez R. (2005):** Inhibition of DNA amplification in marine fish larvae preserved in formalin. *Journal of Plankton Research* **27**(8): 787-792.

**Di Bernardo G., Del Gaudio S., Cammarota M., Galderisi U., Cascino A. and Cipollaro M. (2002):** Enzymatic repair of selected cross-linked homoduplex molecules enhances nuclear gene rescue from Pompeii and Herculaneum remains. *Nucleic Acids Research* **30**(4): e16.

**Diekmann R., Nellen W. and Piatkowski U. (2006):** A multivariate analysis of larval fish and paralarval cephalopod assemblages at Great Meteor Seamount. *Deep-Sea Research I* **53**: 1635-1657.

**Dioh W., Tharreau D., Notteghem J.L., Orbach M. and Lebrun M-H. (2000):** Mapping of avirulence genes in the rice blast fungus, *Magnaporthe grisea*, with RFLP and RAPD markers. *Molecular Plant-Microbe Interactions* **13**(2): 217-227.

**Dioni W. (2002):** TOPICAL TIPS 12: Formalin – Formaldehyde. *Micscape*, **Issue 84:** October 2002. (www.microscopy-uk.org.uk/mag/artoct02/toptips12.html)

**Dolezel J., Bartos J., Voglmayr H. and Greilhuber J. (2003):** Nuclear DNA content and genome size of trout and human. *Cytometry* **51A**(2): 127-128. (from: Gregory T.R. (2005): *Animal Genome Size Database.* http://www.genomesize.com).

**Dorris M. (1999):** Chapter 3 and 4: Reversing the effects on formalin fixation. *Ph.D. Thesis*: *Mode and tempo in the evolution of nematodes.* University of Edinburgh, Scotland, UK. (http://nema.cap.ed.ac.uk/protocols/formalin/formalin.html).

**Douglas M.P. and Rogers S.O. (1998):** DNA damage caused by common cytological fixatives. *Mutation Research* **401**: 77-78.

**Doukakis P., Birstein V.J., DeSalle R., Ludwig A.N., Ludwig A., Machordom A., Almodovar A. and Elvira B. (2000):** Failure to confirm previous identification of two putative museum specimens of the Atlantic sturgeon, *Acipenser sturio*, as the Adriatic sturgeon, *A. naccarii*. *Marine Biology* **136**: 373-377.

**Dowton M. (2004):** Assessing the relative rate of (mitochondrial) genomic change. *Genetics* **167**: 1027-1030.

**Duarte V., Gasparutto D., Jaquinod M. and Cadet J. (2000):** *In vitro* DNA synthesis opposite oxazolone and repair of this DNA damage using modified oligonucleotides. *Nucleic Acids Res* **28:** 1555-1563.

**Dubeau L., Chandler L.A., Gralow J.R., Nichols P.W. and Jones P.A. (1986):** Southern blot analysis of DNA extracted from formalin-fixed pathology specimens. *Cancer Res.* **46**: 2964-2969.

**Dunn C.P. (2003):** Keeping taxonomy based in morphology. *Trends Ecol. Evol.* **18**: 270-271.

**Durand J.D., Persat H. and Bouvet Y. (1999):** Phylogeography and postglacial dispersion of the chub (*Leuciscus cephalus*) in Europe. *Molecular Ecology* **8**: 989-997.

**Echeverrigaray S., Grazziotin G., Grazziotin F. and Agostini G. (2001):** Random amplified polymorphisms between two South American subspecies of rattlesnakes (*Crotalus durissus collilineatus* and *Crotalus durissus terrificus*). *Brazilian Archives of Biology and Technology* 44(3): 313-317.

**Eckerman C.M. (2006):** Formalin-fixed DNA.
(http://www.dmacc.edu/instructors/cmeckerman/formalin/Formalin_fixed_DNA.htm)

**Eckerman C.M. and Walsh E.J. (1997):** Breaking the formalin barrier: Development of molecular techniques for genetic analysis of museum specimens. *Unpublished* (http://www.dmacc.edu/instructors/cmeckerman/formalin/Formalin_fixed_DNA.htm)

**Eckert K.A. and Kunkel T.A. (1990):** High fidelity DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Nucleic Acids Research* **18**(13): 3739-3744.

**Eckert K.A. and Kunkel T.A. (1991):** DNA polymerase fidelity and the polymerase chain reaction. *PCR Methods and Applications* **1**: 17-24.

**Edson S.M., Ross J.P., Coble M.D., Parsons T.J. and Barritt S.M. (2004):** Naming the dead – confronting the realities of rapid identification of degraded skeletal remains. *Forensic Science Review* **16**(1): 64-90.

**Edwards D.D., Deatherage D.E. and Ernsting B.R. (2002):** Random amplified polymorphic DNA analysis of kinship within host-associated populations of the symbiotic water mite *Unionicola foili* (Acari: Unionicolidae). *Experimental and Applied Acarology* **34**(1-2): 67-77.

**Edwards Y.J.K., Elgar G., Clark M.S. and Bishop M.J. (1998):** The identification and characterization of microsatellites in the compact genome of the Japanese pufferfish, *Fugu rubripes*: Perspectives in functional and comparative genomic analyses. *J. Mol. Biol.* **278**: 843-854.

**Ellegren H. (1991):** DNA typing of museum birds. *Nature* **354**: 133.

**Ellsworth D.L., Rittenhouse K.D. and Honeycutt R.L. (1993):** Artifactual variation in Randomly Amplified Polymorphic DNA banding patterns. *BioTechniques* **14**(2): 214-217.

**Enan M.R. (2006):** Application of random amplified polymorphic DNA (RAPD) to detect the genotoxic effect of heavy metals. *Biotechnology and Applied Biochemistry* **43**: 147-154.

**Ender A., Schwenk K., Stadler T., Streit B. and Schierwater B. (1996):** RAPD identification of microsatellites in *Daphnia*. *Molecular Ecology* **5**(3): 437-441.

**Endo K., Noguchi Y., Ueshima R. and Jacobs H.T. (2005):** Novel repetitive structures, deviant protein-encoding sequences and unidentified ORFs in the mitochondrial genome of the brachiopod *Lingula anatina*. *Journal of Molecular Evolution* **61**: 36-53.

**Ennos R.A., French G.C. and Hollingsworth P.M. (2005):** Conserving taxonomic complexity. *TRENDS in Ecology and Evolution* **20** (4): 164-168.

**Etter R.J., Rex M.A., Chase M.C. and Quattro J.M. (1999):** A genetic dimension to deep-sea biodiversity. *Deep-Sea Research I* **46**: 1095-1099.

**Etter R.J. and Mullineaux L. (2001):** Deep-sea communities. pp. 367–393 *in* M. D. Bertness, S. Gaines, and M. Hay, eds. ***Marine community ecology***. Sinauer, Sunderland, MA.

**Etter R.J., Rex M.A., Chase M.R. and Quattro J.M. (2005):** Population differentiation decreases with depth in deep-sea bivalves. *Evolution* **59**(7): 1479-1491.

**Evison M.P., Smillie D.M. and Chamberlain A.T. (1997):** Extraction of single-copy nuclear DNA from forensic specimens with a variety of postmortem histories. *J Forensic Sci.* **42**(6): 1032-1038.

**Fajardo V., Gonzalez I., Lopez-Calleja I., Martin I., Hernandez P.E., Garcia T. and Martin R. (2006):** PCR-RFLP authentication of meats from red deer (*Cervus elaphus*), fallow deer (*Dama dama*), roe deer (*Capreolus capreolus*), cattle (*Bos taurus*), sheep (*Ovis aries*), and goat (*Capra hircus*). *J Agric Food Chem.* **54**(4): 1144-1150.

**Fang S-G., Wan Q-H. and Fujihara N. (2002):** Formalin removal from archival tissue by critical point drying. *BioTechniques* **33**: 604-611.

**Fani R., Damiani G., Di Serio C., Gallori E., Grifoni A. and Bazzicalupo M. (1993):** Use of random amplified polymorphic DNA (RAPD) for generating specific DNA probes for microorganisms. *Mol. Ecol.* **2**: 243-250.

**Faulkner S.W. and Leigh D.A. (1998):** Universal amplification of DNA isolated from small regions of parafin-embedded, formalin-fixed tissue. *BioTechniques* **24**(1): 47-50.

**Fernandez S., Pagotto A.H., Furtado M.M., Katsuyama A.M., Madeira A.M.B.N. and Gruber A. (2003):** A multiplex PCR assay for the simultaneous detection and discrimination of the seven *Eimeria* species that infect domestic fowl. *Parasitology* **127**: 317-325.

**Ferrer I., Armstrong J., Capellari S., Parchi P., Arzberger T., Bell J., Budka H., Strobel T., Giaccone G., Rossi G., Bogdanovic N., Fakai P., Schmitt A., Riederers P., Al-Sarraj S., Ravid R. and Kretzschmar H. (2007):** Effects of formalin fixation, paraffin embedding, and time of storage on DNA preservation in brain tissue: A BrainNet Europe Study. *Brain Pathol* **17**: 297-303.

**Finnerty J.R. and Block B.A. (1992):** Direct sequencing of mitochondrial DNA detects highly divergent haplotypes in blue marlin (*Makaira nigricans*). *Mol. Mar. Biol. Biotechnolog.* **1**: 206-214.

**Finnerty J.R. and Block B.A. (1995):** Evolution of cytochrome b in the Scombroidei (Teleostei): molecular insights into billfish (Istiophoridae and Xipbiidae) relationships. *Fishery Bulletin* **93**: 78-96.

**Fisher C.R. (1995):** Towards an appreciation of hydrothermal-vent animals: their environment, physiological ecology, and tissue stable isotope values. pp. 297-316 *in* S.E. Humphris, R.A. Zierenberg, L.S. Mullineaux and R.E. Thomson, eds. ***Seafloor hydrothermal systems, Physical, Chemical, Biological and Geological interactions.*** Geophysical Monograph No. 91, American Geophysical Union, Washington, DC.

**Fitzhugh K. (2005):** The inferential basis of species hypotheses: the solution to defining the term "species". *Marine Ecology* **26**: 155-165.

**Flaman J.M., Frebourg T., Moreau V., Charbonnier F., Martin C., Ishioka C., Friend S.H. and Iggo R. (1994):** A rapid PCR fidelity assay. *Nucleic Acids Res.* **22**: 3259-3260.

**Floyd R., Abebe E., Papert A. and Blaxter M. (2002):** Molecular barcodes for soil nematode identification. *Molecular Ecology* **11**: 839-850.

**Folmer O., Black M., Hoeh W., Lutz R. and Vrijenhoek R. (1994):** DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotechnol.* **3**(5): 294-299.

**Foissner W., Stoeck T., Schmidt H. and Berger H. (2001):** Biogeographical differences in a common soil ciliate, *Gonostomum affine* (Stein), as revealed by morphological and RAPD-fingerprint analysis. *Acta Protozoologica* **40**: 83-97.

**Fong J.J., Parham J.F., Shi H., Stuart B.L. and Carter R.L. (2007):** A genetic survey of heacily exploited, endangered turtles: caveats on the conservation value of trade animals. *Animal Conservation* **10**: 452-460.

**Forlin L., Livingstone D.R., Magnusson K., Peters L.D., Sole M., Sjolin A. and Granmo A. (1996):** Molecular investigations into pollutant impact on roundnose grenadier (*C. rupestris*) and transplanted common mussel (*M. edulis*) in Skagerrak, the North Sea. *Marine Environmental Research* **42**: 209-212.

**France S.C. and Kocher T.D. (1996):** DNA sequencing of formalin-fixed crustaceans from archival research collections. *Molecular Marine Biology and Biotechnology* **5**(4): 304-313.

**Fredricks D.N. and Relman D.A. (1998):** Improved amplification of microbial DNA from blood cultures by removal of the PCR inhibitor sodium polyanetholesulfonate. *Journal of Clinical Microbiology* **36**(10): 2810-2816.

**Fuchs J., Fjeldsa J. and Pasquet E. (2005):** The use of mitochondrial and nuclear sequence data in assessing the taxonomic status of the endangered Uluguru Bush Shrike *Malaconotus alius*. *Ibis* **147**: 717-724.

**Gach M.H. and Brown W.M. (1997):** Characteristics and distribution of large tandem duplications in brook stickleback (*Culaea inconstans*) mitochondrial DNA. *Genetics* **145**: 383-394.

**Gage J.D. (2004):** Diversity in deep-sea benthic macrofauna: the importance of local ecology, the larger scale, history and the Antarctic. *Deep Sea Research Part II: Topical Studies in Oceanography* **51**: 1689-1708.

**Gage J.D. and Tyler P.A. (1991):** ***Deep-sea biology: a natural history of organisms at the deep-sea floor***. Cambridge University Press, Cambridge, UK.

**Gagliano M., Kowalewsky S. and McCormick M.I. (2006):** An alternative method for the preservation of tropical fish larvae. *Journal of Fish Biology* **68**: 634-639.

**Gaines C.A., Hare M.P., Beck S.E. and Rosenbaum H.C. (2005):** Nuclear markers confirm taxonomic status and relationships among highly endangered and closely related right whale species. *Proc. R. Soc. B* **272**: 533-542.

**Galtier N., Enard D., Radondy Y., Bazin E. and Belkhir K. (2006):** Mutation hot spots in mammalian mitochondrial DNA. *Genome Research* **16**: 215-222.

**Gamache I., Jaramillo-Correa J.P., Payette S. and Bousquet J. (2003):** Diverging patterns of mitochondrial and nuclear DNA diversity in subarctic black spruce: imprint of a founder effect associated with postglacial colonization. *Molecular Ecology* **12**(4): 891-901.

**Garcia D.K., Faggart M.A., Rhoades L. and Alcivar-Warren A.A. (1994):** Genetic diversity of cultured *Penaeus vannamei* shrimp using three molecular techniques. *Molecular Marine Biology and Biotechnology* **3**(5): 270-280.

**Garrido-Ramos M.A., Soriguer M.C., Herran R de la, Jamilena M., Ruiy-Rejon C., Domezain A., Hernando J.A. and Ruiy-Rejon M. (1997):** Morphometric and genetic analysis as proof of the existence of two sturgeon species in the Guadalquivir river. *Marine Biology* **129**: 33-39.

**Gebicki S. and Gebicki J.M. (1999):** Crosslinking of DNA and proteins induced by protein hydroperoxides. *Biochem. J.* **338**: 629-636.

**Geller J.B., Carlton J.T. and Powers D.A. (1993):** Interspecific and intrapopulation variation in mitochondrial ribosomal DNA sequences of *Mytilus* spp. (Bivalvia: Mollusca). *Mol. mar. Biol. Biotech.* **2**: 44-50.

**Genner M.J., Sims D.W., Wearmouth V.J., Southall E.J., Southward A.J., Henderson P.A. and Hawkins S.J. (2004):** Regional climatic warming drives long-term community changes of British marine fish. *Proc. R. Soc. Lond. B* **271**: 655-661.

**George M. Jr., Puentes L.A. and Ryder O.A. (1993):** Genetic differences between subspecies of white rhinoceros. ***International studbook of African rhinos.*** Berlin Zoological Garden, Berlin, East Germany.

**Gerloff U., Schlotterer C., Rassmann K., Rambold I., Hohmann G., Fruth B. and Tautz D. (1995):** Amplification of hypervariable simple sequence repeats (microsatellites) from excremental DNA of wild living bonobos (*Pan paniscus*). *Molecular Ecology* **4**: 515-518.

**Gershwin L. (2002):** Curating Medusae. *Electronic internet document* (http://www.medusozoa.com/curating.html)

**Ghany A.G.A.A. and Zaki E.A. (2003):** DNA sequences of RAPD fragments in the Egyptian cotton *Gossypium barbadense*. *African Journal of Biotechnology* **2**(5): 129-132.

**Giannella C., Zito F.A., Colonna F., Paradiso A., Marzullo F., Alaibac M. and Schittulli F. (1997):** Comparison of formalin, ethanol, and Histochoice fixation on the PCR amplification from paraffin-embedded breast cancer tissue. *Eur J Clin Chem Clin Biochem* **35**(8): 633-635.

**Gibson G. and Muse S. V. (2002):** ***A Primer of Genome Science***. Sinauer Associates, Inc. Publisheres, Sunderland, Massachusetts, U.S.A.

**Gilbert M.T.P. (2003a):** An assessment of the use of human samples in ancient DNA studies. *Ph.D. Thesis.* Henry Wellcome Ancient Biomolecules Centre, Department of Zoology, University of Oxford, U.K.

**Gilbert M.T.P. (2003b):** Distribution patterns of postmortem damage in human mitochondrial DNA. *Am. J. Hum. Genet.* **72**: 32-47.

**Gilbert M.T.P., Willerslev E., Hansen A.J., Barnes I., Rudbeck L., Lynnerup N. and Cooper A. (2003):** Distribution patterns of postmortem damage in human mitochondrial DNA. *American Journal of Human Genetics* **72**(7): 625-634.

**Gilbert M.T.P., Bandelt H-J., Hofreiter M. and Barnes I. (2005a):** Assessing ancient DNA studies. *TRENDS in Ecology and Evolution* **20**(10): 541-544.

**Gilbert M.T.P., Shapiro B., Drummond A. and Cooper A. (2005b):** Post-mortem DNA damage hotspots in Bison (*Bison bison*) provide evidence for both damage and mutational hotspots in human mitochondrial DNA. *Journal of Archaeological Science* **32**: 1053-1060.

**Gilbert M.T.P., Binladen J., Miller W., Wiuf C., Willerslev E., Poinar H., Carlson J.E., Leebens-Mack J.H. and Schuster S.C. (2007a):** Recharacterization of ancient DNA miscoding lesions: insights in the era of sequencing-by-synthesis. *Nucleic Acids Research* **35**(1): 1-10.

**Gilbert M.T.P., Haselkorn T., Bunce M., Sanchez J.J., Lucas S.B., Jewell L.D., Van Marck E. and Worobey M. (2007b):** The isolation of nucleic acids from fixed, paraffin-embedded tissues – which methods are useful when? *PLoS ONE* **2**(6): e537.

**Gilbert M.T.P., Sanchez J.J., Haselkorn T., Jewell L.D., Lucas S.B., Van Marck E., Borsting C., Morling N. and Worobey M. (2007c):** Multiplex PCR with minisequencing as an effective high-throughput SNP typing method for formalin-fixed tissue. *Electrophoresis* **28**: 2361-2367.

**Gilbert M.T.P., Moore W., Melchior L. and Worobey M. (2007d):** DNA extraction from dry museum beetles without conferring external morphological damage. *PLoS ONE* **2**(3): e272.

**Giles R.E., Blanc H., Cann H.M. and Wallace D.C. (1980):** Maternal inheritance of human mitochondrial DNA. *Proc. Natl. Acad. Sci. U.S.A.* **77**: 6715-6719.

**Gill P. (2001):** Application of low copy number DNA profiling. *Croatian Medical Journal* **42**(3): 229-232.

**Gissi C., Reyes A., Pesole G. and Saccone C. (2000):** Lineage-specific evolutionary rate in mammalian mtDNA. *Mol. Biol. Evol.* **17**(7): 1022-1031.

**Gjetvaj B., Cook D.I. and Zouros E. (1992):** Repeated sequences and large-scale size variation of mitochondrial DNA: A common feature among scallops (Bivalvia: Pectinidae). *Mol. Biol. Evol.* **9**(1): 106-124.

**Glover A.G. and Smith C.R. (2003):** The deep-sea floor ecosystem: current status and prospects of anthropogenic change by the year 2025. *Environmental Conservation* **30**: 219-241.

**Godhe A., Anderson D.M. and Rehnstam-Holm A-S. (2002):** PCR amplification of microalgal DNA for sequencing and species identification: studies on fixatives and algal growth stages. *Harmful Algae* **1**: 375-382.

**Goelz S.E., Hamilton S.R. and Vogelstein B. (1985):** Purification of DNA from formaldehyde fixed and paraffin embedded human tissue. *Biochem Biophys Res Commun* **130**(1): 118-126.

**Goldstein P.Z. and DeSalle R. (2003):** Calibrating phylogenetic species formation in a threatened insect using DNA from historical specimens. *Molecular Ecology* **12**: 1993-1998.

**Golenberg E.M., Bickel A. and Weihs P. (1996):** Effect of highly fragmented DNA on PCR. *Nucleic Acids Research* **24**(24): 5026-5033.

**Gomes C., Dales R.B.G. and Oxenford H.A. (1998):** The application of RAPD markers in stock discrimination of the four-wing flyingfish, *Hirundichthys affinis* in the central western Atlantic. *Molecular Ecology* **7**: 1029-1039.

**Goodman M.F. (2002):** Error-prone repair DNA polymerases in prokaryotes and eukaryotes. *Annu. Rev. Biochem.* **71**: 17-50.

**Goodman M.F. and Tippin B. (2000):** Sloppier copier DNA polymerases involved in genome repair. *Curr Opin Genet Dev* **10**: 162-168.

**Gordon J.D.M. (2001):** Deep-water fish and fisheries: introduction. *Fisheries Research* **51**: 105-111.

**Gouin N., Grandjean F., Bouchon D., Reynolds J.D. and Souty-Grosset C. (2001):** Population genetic structure of the endanered freshwater crayfish *Austropotamobius pallipes*, assessed using RAPD markers. *Heredity* **87**: 80-87.

**Graham C.H., Ferrier S., Huettman F., Moritz C. and Peterson A.T. (2004):** New developments in museum-based informatics and applications in biodiversity analysis. *TRENDS in Ecology and Evolution* **19**(9): 497-503.

**Grassle J.F. (1977):** Slow recolonisation of deep-sea sediment. *Nature, London* **265**: 618-619.

**Gray M.W., Lang B.F., Cedergren R., Golding G.B., Lemieux C., Sankoff D., Turmel M., Brossard N., Delage E., Littlejohn T.G., Plante I., Rioux P., Saint-Louis D., Zhu Y. and Burger G. (1998):** Genome structure and gene content in protist mitochondrial DNAs. *Nucleic Acids Research* **24**(4): 865-878.

**Grazziotin F. and Echeverrigaray S. (2005):** Genetic relationships among species of the genus *Bothrops* based on RAPD markers. *Brazilian Archives of Biology and Technology* **48**(3): 359-365.

**Greagg M.A., Fogg M.J., Panayotou G., Evans S.J., Connolly B.A. and Pearl L.H. (1999):** A read-ahead function in archaeal DNA polymerases detects promutagenic template-strand uracil. *Proc. Natl. Acad. Sci. USA* **96**: 9045-9050.

**Greenwood A.D. (2001):** Mammoth biology: Biomolecules, phylogeny, numts, nuclear DNA, and the biology of an extinct species. *Ancient Biomolecules* **3**: 255-266.

**Greer C.E., Peterson S.L., Kiviat N.B. and Manos M.M. (1991):** PCR amplification from paraffin-embedded tissues-effects of fixative and fixation time. *American Journal of Clinical Pathology* **95**(2): 117-124.

**Grosberg R.K., Levitan D.R. and Cameron B.B. (1996):** Characterization of genetic structure and genealogies using RAPD-PCR markers: A random primer for the novice and nervous. Pages 67-100 *in* Ferraris J.D. and Palumbi S.R., eds. ***Molecular Zoology: Advances, Strategies, and Protocols.*** Wiley-Liss, Inc., New York, USA.

**Grothues D., Cantor C.R. and Smith C.L. (1993):** PCR amplification of megabase DNA with tagged random primers (T-PCR). *Nucleic Acids Research* **21**(5): 1321-1322.

**Guerra-Garcia J.M., Redondo-Gomez S., Espina A.G., Castillo J.M., Luque T., Garcia-Gomez J.C. and Figueroa M.E. (2006):** *Caprella penantis* Leach, 1814 and *Caprella dilatata* Kroyer, 1843 (Crustacea: Amphipoda) from the Strait of Gibraltar: a molecular approach to explore intra- and interspecific variation. *Marine Biology Research* **2**: 100-108.

**Guha S. and Kashyap V.K. (2006):** Molecular identification of lizard by RAPD & FINS of mitochondrial 16S rRNA gene. *Legal Medicine* **8**: 5-10.

**Guo X., Liu S. and Liu Y. (2006):** Evidence for recombination of mitochondrial DNA in triploid crucian carp. *Genetics* **172**: 1745-1749.

**Gurdebeke S. and Maelfait J-P. (2002):** Pitfall trapping in population genetics studies: finding the right "solution". *The Journal of Arachnology* **30**: 255-261.

**Gyllensten U. (1985):** The genetic structure of fish: differences in the intraspecific distribution of biochemical variation between marine, anadromous, and freshwater species. *J. Fish Biol.* **26**: 691-699.

**Gyllensten U., Wharton D., Josefsson A. and Wilson A.C. (1991):** Paternal inheritance of mitochondrial DNA in mice. *Nature* **352**: 255-257.

**Hadrys H., Balick M. and Schierwater B. (1992):** Application of random amplified polymorphic DNA (RAPD) in molecular ecology. *Molecular Ecology* **1**: 55-63.

**Hadrys H. and Siva-Jothy M.T. (1994):** Unravelling the components that underlie insect reproductive traits using a simple molecular approach. Pages 75-90 *in* Schierwater B., Streit B., Wagner G.P. and DeSalle R., editors. ***Molecular ecology and evolution: Approaches and applications***. Birkhauser Verlag, Basel, Boston, Berlin

**Haedrich R.L. and Barnes S.M. (1997):** Changes over time of the size structure in an exploited shelf fish community. *Fisheries Research* **31**(3): 229-239.

**Haedrich R.L., Merrett N.R. and O'Dea N.R. (2001):** Can ecological knowledge catch up with deep-water fishing? A North Atlantic perspective. *Fisheries Research* **51**: 113-122.

**Hagelberg E., Sykes B. and Hedges R. (1989):** Ancient bone DNA amplified. *Nature* **342**: 485.

**Haglund W.D. and Sorg M.H. (1997):** Advances in *forensic taphonomy: Method, theory, and archaeological perspectives*. CRC Press, Boca Raton, London, New York, Washington.

**Haglund W.D. and Sorg M.H. (2002):** *Forensic taphonomy: The postmortem fate of human remains*. CRC Press, Boca Raton, London, New York, Washington.

**Hajibabaei M., deWaard J.R., Ivanova N.V., Ratnasingham S., Dooh R.T., Kirk S.L., Mackie P.M. and Hebert P.D.N. (2005):** Critical factors for assembling a high volume of DNA barcodes. *Phil. Trans. R. Soc. B* **360**(1462): 1959-1967.

**Hajibabaei M., Smith M.A., Janzen D.H., Rodriguez J.J., Whitfield J.B. and Hebert P.D.N. (2006):** A minimalist barcode can identify a specimen whose DNA is degraded. *Molecular Ecology Notes* **6**(4): 959-964.

**Hale W.G., Margham J.P. and Saunders V.A. (1995):** *Collins Dictionary - Biology*. HarperCollins Publishers, U.K.

**Hallden C., Hansen M., Nilsson N.-O., Hjerdin A. and Sall T. (1996):** Competition as a source of errors in RAPD analysis. *Theoretical and Applied Genetics* **93**(8): 1185-1192

**Hammond J.B.W. (1995):** *An introduction to gene cloning and analysis*. Course book, North East Surrey College of Technology (NESCOT), Ewell East, Surrey, UK.

**Hammond J.B.W. (1997):** *An introduction to gene cloning and analysis*. Updated course book, North East Surrey College of Technology (NESCOT), Ewell East, Surrey, UK.

**Hammond J.B.W., Spanswick G. and Mawn J.A. (1996):** Extraction of DNA from preserved animal specimens for use in randomly amplified polymorphic DNA analysis. *Analytical Biochemistry* **240**: 298-300.

**Hanni C., Lauder V., Sakka M., Begue A. and Stehelin D. (1990):** Amplification of mitochondrial DNA fragments from ancient human teeth and bones. *C.R. Acad. Sci. III* **310**: 365-370.

**Hanni C., Brousseau T., Laudet V. and Stehelin D. (1995):** Isopropanol precipitation removes PCR inhibitors from ancient bone extracts. *Nucleic Acids Research* **23**(5): 881-882.

**Hannington M.D., Jonasson I.R., Herzig P.M. and Petersen S. (1995):** Physical and chemical processes of seafloor mineralization at mid-ocean Ridges. pp. 115-157 *in:* S.E. Humphris, R.A. Zierenberg, L.S. Mullineaux and R.E. Thomson, eds. ***Seafloor hydrothermal systems, Physical, Chemical, Biological and Geological interactions.*** Geophysical Monograph No. 91, American Geophysical Union, Washington, DC.

**Hansen A.J., Willerslev E., Wiuf C., Mourier T. and Arctande P. (2001):** Statistical evidence for miscoding lesions in ancient DNA templates. *Mol. Biol. Evol.* **18**(2): 262-265.

**Hansen A.J., Mitchell D.L., Wiuf C., Paniker L., Brand T.B., Binladen J., Gilichinsky D.A., Ronn R. and Willerslev E. (2006):** Crosslinks rather than strand breaks determine access to ancient DNA sequences from frozen sediments. *Genetics* **173**: 1175-1179.

**Hansen M., Hallden C. and Sall T. (1998):** Error rates and polymorphism frequencies for three RAPD protocols. *Plant Molecular Biology Reporter* **16**: 139-146.

**Harper G.L., Maclean N. and Goulson D. (2006):** Analysis of museum specimens suggests extreme genetic drift in the adonis blue butterfly (*Polyommatus bellargus*). *Biological Journal of the Linnean Society* **88**: 447-452.

**Harr B. and Schlotterer C. (2000):** Long microsatellite alleles in *Drosophila melanogaster* have a downward mutation bias and short persistence times, which cause their genome-wide underrepresentation. *Genetics* **155**: 1213-1220.

**Harris P.M., Roe K.J. and Mayden R.L. (2005):** A mitochondrial DNA perspective on the molecular systematics of the sunfish genus *Lepomis* (Actinopterygii: Centrarchidae). *Copeia* **2005**(2): 340-346.

**Hasbun C.R., Gomez A., Kohler G. and Lunt D.H. (2005):** Mitochondrial DNA phylogeography of the Mesoamerican spiny-tailed lizards (*Ctenosaura quinquecarinata* complex): historical biogeography, species status and conservation. *Molecular Ecology* **14**: 3095-3107.

**Haselkorn R. and Doty P. (1961):** The reaction of formaldehyde with polynucleotides. *J. Biol. Chem.* **236**: 2738-2748.

**Hassanin A. (2002):** Ancient specimens and DNA contamination: a case study from the 12S rRNA gene sequence of the "Linh Duong" bovid (*Pseudonovibos spiralis*). *Naturwissenschaften* **89**: 107-110.

**Hauser L., Carvalho G.R. and Pitchier T.J. (1998):** Genetic population structure in the Lake Tanganyika sardine *Limnothrissa miodon*. *Journal of Fish Biology* **53**(Supplement A): 413-429.

**Haussermann V. (2004):** Identification and taxonomy of soft-bodied hexacorals exemplified by Chilean sea anemones; including guidelines for sampling, preservation and examination. *J. Mar. Biol. Ass. U.K.* **84**: 931-936.

**Hayashi J.-I., Tagashira Y. and Yoshida M.C. (1985):** Absence of extensive recombination between inter- and intraspecies mitochondrial DNA in mammalian cells. *Exp. Cell. Res.* **160**: 387-395.

**Hebert P.D.N., Cywinska A., Ball S.L. and deWaard J. R. (2003):** Biological identifications through DNA barcodes. *Proc. R. Soc. Lond. Ser. B* **270**: 313–321.

**Hebert P.D.N., Penton E.H., Burns J.M., Janzen D.H. and Hallwachs W. (2004):** Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *PNAS* **101**(41): 14812-14817.

**Hebsgaard M.B., Phillips M.J. and Willerslev E. (2005):** Geologically ancient DNA: fact or artefact? *TRENDS in Microbiology* **13**(5): 212-220.

**Helfenbein K.G., Fourcade H.M., Vanjani R.G. and Boore J.L. (2004):** The mitochondrial genome of *Paraspadella gotoi* is highly reduced and reveals that chaetognaths are a sister group to protostomes. *PNAS* **101**(29): 10639-10643.

**Hellberg M.E. (2006):** No variation and low synonymous substitution rates in coral mtDNA despite high nuclear variation. *BMC Evolutionary Biology* **6**: 24.

**Helly J.J. and Levin L.A. (2004):** Global distribution of naturally occurring marine hypoxia on continental margins. *Deep-Sea Research I* **51**: 1159-1168.

**Herniou E.A., Pearce A.C. and Littlewood D.T.J. (1998):** Vintage helminths yield valuable molecules. *Parasitology Today* **14**(7): 289-292.

**Herrmann B. and Hummel S. (1994):** ***Ancient DNA: recovery and analysis of genetic material from paleontological, archeological, museum, medical, and forensic specimens.*** Springer-Verlag, New York.

**Hey J., Waples R.S., Arnold M.L., Butlin R.K. and Harrison R.G. (2003):** Understanding and confronting species uncertainty in biology and conservation. *TRENDS in Ecology and Evolution* **18**(11): 597-603.

**Higgins D.G. and Sharp P.M. (1988):** CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene* **73**: 237-244.

**Higuchi R., Dollinger G., Walsh P.S. and Griffith R. (1992):** Simultaneous amplification and detection of specific DNA sequences. *Biotechnology* **10**: 413-417.

**Higuchi R., Fockler C., Dollinger G. and Watson R. (1993):** Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. *Biotechnology* **11**: 1026-1030.

**Hill W.G. and Weir B.S. (2004):** Moment estimation of population diversity and genetic distance from data on recessive markers. *Molecular Ecology* **13**: 895-908.

**Hillis D.M. and Dixon M.T. (1991):** Ribosomal DNA: Molecular evolution and phylogenetic inference. *The Quarterly Review of Biology* **66**(4): 411-453.

**Hillis D.M., Moritz C. and Mable B.K. (1996):** ***Molecular systematics***, Second edition. Sinauer Associates Inc., Sunderland, Massachusetts, USA.

**Hoarau G., Holla S., Lescasse R., Stam W.T. and Olsen J.L. (2002):** Heteroplasmy and evidence for recombination in the mitochondrial control region of the flatfish *Platichthys flesus*. *Mol Biol Evol* **19**(12): 2261-2264.

**Hoarau G., Piquet A.M.-T., van der Veer H.W., Rijnsdorp A.D., Stam W.T. and Olsen J.L. (2004):** Population structure of plaice (*Pleuronectes platessa* L.) in northern Europe: a comparison of resolving power between microsatellites and mitochondrial DNA data. *Journal of Sea Research* **51**(3-4): 183-190.

**Hoffman J.I. and Amos W. (2005):** Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. *Molecular Ecology* **14**: 599-612.

**Hoff-Olsen P., Mevag B., Staalstrom E., Hovde B., Egeland T. and Olaisen B. (1999):** Extraction of DNA from decomposed human tissue: An evaluation of five extraction methods for short tandem repeat typing. *Forensic Science International* **105**: 171-183.

**Hofmann G.E., Burnaford J.L. and Fielman K.T. (2005):** Genomic-fueled aproaches to current challenges in marine ecology. *TRENDS in Ecology and Evolution* **20**(6): 305-311.

**Hofreiter M., Serre D., Poinar H.N., Kuch M. and Paabo S. (2001a):** Ancient DNA. *Nature Reviews/ Genetics* **2**: 353-359.

**Hofreiter M., Jaenicke V., Serre D., von Haeseler A. and Paabo S. (2001b):** DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. *Nucleic Acids Research* **29**(23): 4793-4799.

**Hogan Z.S. and May B.P. (2002):** Twenty-seven new microsatellites for the migratory Asian catfish family Pangasiidae. *Molecular Ecology Notes* **2**: 38-41.

**Holt R.D., Gomulkiewicz R. and Barfield M. (2003):** The phenomenology of niche evolution via quantitative traits in a "black-hole" sink. *Proc R. Soc. Lond. B* **270**: 215-224.

**Hoorfar J., Wolffs P. and Radstrom P. (2004):** Diagnostic PCR: validation and sample preparation are two sides of the same coin. *APMIS* **112**: 808-814.

**Hoss M., Jaruga P., Zastawny T.H., Dizdaroglu M. and Paabo S, (1996):** DNA damage and DNA sequence retrieval from ancient tissues. *Nucleic Acids Research* **24**(7): 1304-1307.

**Huang G., Lu H., Hao A., Ng D.C.H., Ponniah S., Guo K., Lufei C., Zeng Q. and Cao X. (2004):** GRIM-19, a cell death regulatory protein, is essential for assembly and function of mitochondrial complex I. *Molecular and Cellular Biology* **24**(19): 8447-8456.

**Huber J.T. (1998):** The importance of voucher specimens, with practical guidelines for preserving specimens of the major invertebrate phyla for identification. *Journal of Natural History* **32**: 367-385.

**Hughes C.E., Eastwood R.J. and Bailey D. (2006):** From famine to feast? Selecting nuclear DNA sequence loci for plant species-level phylogeny reconstruction. *Phil. Trans. R. Soc. B* **361**: 211-225.

**Hulley P. A. (1986):** A taxonomic review of the lanternfish genus *Triphoturus* Fraser-Brunner, 1949 (Myctophidae, Osteichthyes). *Annals of the South African Museum* **97**: 71–95.

**Hummel S. (2003):** ***Ancient DNA Typing: methods, strategies and applications.*** Springer-Verlag, Berlin Heidelberg.

**Hunt G.J. and Page R.E. Jr. (1992):** Patterns of inheritance with RAPD molecular markers reveal novel types of polymorphism in the honey bee. *Theor. Appl. Gen.* **85**: 15-20.

**Hurwood D.A., Hughes J.M., Bunn S.E. and Cleary C. (2003):** Population structure in the freshwater shrimp (*Paratya australiensis*) inferred from allozymes and mitochondrial DNA. *Heredity* **90**: 64-70.

**Huynen L., Millar C.D., Scofield R.P. and Lambert D.M. (2003):** Nuclear DNA sequences detect species limits in ancient moa. *Nature* **425**: 175-178.

**Huynen M.A. and Bork P. (1998):** Measuring genome evolution. *Proc. Natl. Acad. Sci. USA* **95**: 5849-5856.

**Inadome Y. and Noguchi M. (2003):** Selection of higher molecular weight genomic DNA for molecular diagnosis from formalin-fixed material. *Diagnostic Molecular Pathology* **12**(4): 231-236.

**Iniguez A.M., Araujo A., Ferreira L.F. and Vicente A.C.P. (2003):** Analysis of ancient DNA from Coprolites: a perspective with random amplified polymorphic DNA-polymerase chain reaction approach. *Mem Inst Oswaldo Cruz, Rio de Janeiro* **98**(Suppl. I): 63-65.

**Inoue J.G., Miya M., Tsukamoto K. and Nishida M. (2003):** Evolution of the deep-sea gulper eel mitochondrial genomes: Large-scale gene rearrangements originated within the eels. *Mol. Biol. Evol.* **20**(11): 1917-1924.

**Irestedt M., Ohlson J.I., Zuccon D., Kallersjo M. and Ericson P.G.P. (2006):** Nuclear DNA from old collections of avian study skins reveals the evolutionary history of the Old World suboscines (Aves, Passeriformes). *Zoologica Scripta* **35**: 567-580.

**Isabel N., Beaulieu J., Theriault P. and Bousquet J. (1999):** Direct evidence for biased gene diversity estimates from dominant random amplified polymorphic DNA (RAPD) fingerprints. *Molecular Ecology* **8**: 477-483.

**Isoda K., Shiraishi S., Watanabe S. and Kitamura K. (2000):** Molecular evidence of natural hybridization between *Abies veitchii* and *A. homolepis* (Pinaceae) revealed by chloroplast, mitochondrial and nuclear DNA markers. *Molecular Ecology* **9**: 1965-1974.

**Ivanova N.V., deWaard J.R. and Hebert P.D.N. (2006):** An inexpensive, automation-friendly protocol for recovering high-quality DNA. *Molecular Ecology Notes* **6**(4): 998-1002.

**Ivanova N.V., Zemlak T.S., Hanner R.H. and Hebert P.D.N. (2007):** Universal primer cocktails for fish DNA barcoding. *Molecular Ecology Notes* **7**(4): 544-548.

**Iwamoto T. (1970):** Macrourid fishes of the Gulf of Guinea. The R/V Pillsbury deep-sea biological expedition to the Gulf of Guinea, 1964-65. *Stud. Tropical Oceanogr.* (Miami) No. 4 (pt 2): 316-431.

**Iwamoto T. and Merrett N.R. (1997):** Pisces, Gadiformes: taxonomy of grenadiers of the New Caledonian region, southwest Pacific. ***In:*** A. Crosnier (ed.), Resultats des Campagnes MUSORSTOM, Volume 18. *Memoires du Museum nationale d'Histoire naturelle*, **176:** 473-570.

**Iwamoto T., McMillan P. and Shcherbachev Y.N. (1999):** A new grenadier, genus *Caelorinchus*, from Australia and New Zealand (Pisces, Gadiformes, Macrouridae). *New Zealand Journal of Marine and Freshwater Research* **33**: 49-54.

**Iyengar A., Piyapattanakorn S., Heipel D.A., Stone D.M., Howell B.R., Child A.R. and Maclean N. (2000):** A suite of highly polymorphic microsatellite markers in turbot (*Scophthalmus maximus* L.) with potential for use across several flatfish species. *Molecular Ecology* **9**: 368-371.

**Jackson D.P., Lewis F.A., Taylor G.R., Boylston A.W. and Quirke P. (1990):** Tissue extraction of DNA and RNA and analysis by the polymerase chain reaction. *Journal of Clinical Pathology* **43**: 499-504.

**Jacobson K.M., Miller O.K. Jr. and Turner B.J. (1993):** Randomly amplified polymorphic DNA markers are superior to somatic incompatibility tests for discrimination genotypes in natural populations of the ectomycorrhizal fungus *Suillus ganulatus*. *Proc. Natl. Acad. Sci. USA* **90**: 9159-9163.

**Jakobsdottir K.B., Jorundsdottir I.D., Skirnisdottir S., Hjorleifsdottir S., Hreggvidsson G.O., Danielsdottir A.K. and Pampoulie C. (2006):** Nine new polymorphic microsatellite loci for the amplification of archived otolith DNA of Atlantic cod, *Gadus morhua* L. *Molecular Ecology Notes* **6**: 337-339.

**James M.J., Yabsley M.J., Pung O.J. and Grijalva M.J. (2002):** Amplification of *Trypanosoma cruzi*-specific DNA sequences in formalin-fixed raccoon tissues using polymerase chain reaction. *J. Parasitol.* **88**(5): 989-993.

**Janzen D.H., Hajibabaei M., Burns J.M., Hallwachs W., Remigio E. and Hebert P.D.N. (2005):** Wedding biodiversity inventory of a large and complex Lepidoptera fauna with DNA barcoding. *Philosophical Transactions: Biological Sciences* **360**: 1835-1845

**Jaremko M., Justenhoven C., Abraham B.K., Schroth W., Fritz P., Brod S., Vollmert C., Illig T. and Brauch H. (2005):** MALDI-TOF MS and TaqMan assisted SNP genotyping of DNA isolated from formalin-fixed and paraffin-embedded tissues (FFPET). *Human Mutation* **25**: 232-238.

**Jeffreys A.J., Wilson V. and Thein S.L. (1985):** Individual-specific "fingerprints" of human DNA. *Nature* **316**: 76-79.

**Jennings S. and Kaiser M.J. (1998):** The effects of fishing on marine ecosystems. *Advances in Marine Biology* **34**: 201-352.

**Ji Y-J., Zhang D-X. and He L-J. (2003):** Evolutionary conservation and versatility of a new set of primers for amplifying the ribosomal internal transcribed spacer regions in insects and other invertebrates. *Molecular Ecology Notes* **3**: 581-585.

**Johansen M.E., Muller J.G., Xu X. and Burrows C.J. (2005):** Oxidatively induced DNA-protein cross-linking between single-stranded binding protein and oligodeoxynucleotides containing 8-Oxo-7,8-dihydro-2`-deoxyguanosine. *Biochemistry* **44**: 5660-5671.

**Johns G.C. and Avise J.C. (1998):** A comparative summary of genetic distances in the vertebrates from the mitochondrial cytochrome *b* gene. *Mol. Biol. Evol.* **15**(11): 1481-1490.

**Johnson A.A. and Johnson K.A. (2001):** Exonuclease proofreading by human mitochondrial DNA polymerase. *The Journal of Biological Chemistry* **276**(41): 38097-38107.

**Jollivet D., Desbruyeres D., Ladrat C. and Laubier L. (1995):** Evidence for differences in the allozyme thermostability of deep-sea hydrothermal vent polychaetes Alvinellidae: a possible selection by habitat. *Marine Ecology Progress Series* **123**: 125-136.

**Jones C.J., Edwards K.J., Castaglione S., Winfield M.O., Sala F., van de Wiel C., Bredemeijer G., Vosman B., Matthes M., Daly A., Brettschneider R., Bettini P., Buiatti M., Maestri E., Malcevschi A., Marmiroli N., Aert R., Volckaert G., Rueda J., Linacero R., Vazquez A. and Karp A. (1997):** Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Molecular Breeding* **3:** 381-390.

**Jovanovic L., Delahunt B., McIver B., Eberhardt N.L. and Grebe S.K.G. (2003):** Optimising restriction enzyme cleavage of DNA derived from archival histopathological samples: an improved HUMARA assay. *Pathology* **35:** 70-74.

**Junqueira A.C., Lessinger A.C. and Azeredo-Espin A.M. (2002):** Methods for the recovery of mitochondrial DNA sequences from museum specimens of myiasis-causing flies. *Medical and Veterinary Entomology* **16:** 39-45.

**Kajander O.A., Karhunen P.J., Holt I.J. and Jacobs H.T. (2001):** Prominent mitochondrial DNA recombination intermediates in human heart muscle. *EMBO Rep* **2:** 1007-1012.

**Kalmar T., Bachrati C.Z., Marcsik A. and Rasko I. (2000):** A simple and efficient method for PCR amplifiable DNA extraction from ancient bones. *Nucleic Acids Research* **28**(12): *e*67.

**Kambhapati S., Black IV W.C. and Rai K.S. (1992):** Random amplified polymorphic DNA of mosquito species and populations (Diptera: Culicidae): techniques, statistical analysis, and applications. *Journal of Medical Entomology* **29**(6): 939-945.

**Karaiskou N., Triantafyllidis A., Alvarez P., Lopes P., Garcia-Vazquez E. and Triantaphyllidis C. (2007):** Horse mackerel egg identification using DNA methodology. *Marine Ecology* **28**: 429-434.

**Karlsen F., Kalantari M., Chitemerere M., Johansson B. and Hagmar B. (1994):** Modifications of human and viral deoxyribonucleic acid by formaldehyde fixation. *Lab Invest* **71:** 604-611.

**Kaufman B.A., Newman S.M., Hallberg R.L., Slaughter C.A., Perlman P.S. and Butow R.A. (2000):** *In organello* formaldehyde crosslinking of proteins to mtDNA: Identification of bifunctional proteins. *PNAS* **97**(14): 7772-7777.

**Kaur M. and Makrigiorgos G.M. (2003):** Novel amplification of DNA in a hairpin structure: towards a radical elimination of PCR errors from amplified DNA. *Nucleic Acids Research* **31**(6): e26.

**Kearney M. and Stuart B.L. (2004):** Repeated evolution of limblessness and digging heads in worm lizards revealed by DNA from old bones. *Proc. R. Soc. Lond. B* **271:** 1677-1683.

**Kendall V.J. and Haedrich R.L. (2006):** Species richness in Atlantic deep-sea fishes assessed in terms of the mid-domain effect and Rapoport's rule. *Deep-Sea Research I* **53:** 506-515.

**Keohavong P. and Thilly W.G. (1989):** Fidelity of DNA polymerases in DNA amplification. *Proc. Natl. Acad. Sci. USA* **86**: 9253-9257.

**Key J.M. (2002):** A review of current knowledge describing New Zeland's deepwater benthic biodiversity. *Marine Biodiversity Biosecurity Report No. 1.*

**Kiernan J.A. (2000):** Formaldehyde, formalin, paraformaldehyde and glutaraldehyde: What they are and what they do. *Microscopy Today* **00-1**: 8-12.

**Kiesslich J., Radacher M., Neuhuber F., Meyer H.J. and Zeller K.W. (2002):** On the use of nitrocellulose membranes for dialysis-mediated purification of ancient DNA from human bone and teeth extracts. *Ancient Biomolecules* **4**(2): 79-87.

**Kilpatrick C.W. (2002):** Noncryogenic preservation of mammalian tissues for DNA extraction: An assessment of storage methods. *Biochemical Genetics* **40**: 53-62.

**Kimmel M., Chakraborty R., King J.P., Bamshad M., Watkins W.S. and Jorde L.B. (1998):** Satures of population expansion in macrosatellite repeat data. *Genetics* **148**: 1921-1930.

**King J.R. and Porter S.D. (2004):** Recommendations on the use of alcohols for preservation of ant specimens (Hymenoptera, Formicidae). *Insect. Soc.* **51**: 197-202.

**King R.A., Read D.S., Traugott M. and Symondson W.O.C. (2008):** Molecular analysis of predation: a review of best practice for DNA-based approaches. *Marine Ecology* **17**: 947-963.

**Kirby R.R. and Reid P.C. (2001):** PCR from the CPR offers a historical perspective on marine population ecology. *J. Mar. Biol. Ass. U.K.* **81**: 539-540.

**Klanten S.O., van Herwerden L. and Choat J.H. (2003):** Acquiring reef fish DNA sequences from formalin-fixed museum specimens. *Bulletin of Marine Science* **73**(3): 771-776.

**Klinbunga S., Ampayup P., Tassanakajon A., Jarayabhand P. and Yoosukh W. (2001):** Genetic diversity and molecular markers of cupped oysters (Genera *Crassostrea, Saccostrea,* and *Striostrea*) in Thailand revealed by randomly amplified polymorphic DNA analysis. *Marine Biotechnology* **3**: 133-144.

**Knowlton N., Weight L.A., Solorzano L.A., Mills D.K. and Bermingham E. (1993):** Divergence in proteins, mitochondrial DNA, and reproductive compatibility across the isthmus of Panama. *Science* **260**: 1629-1632.

**Kocher T.D., Thomas W.K., Meyer A., Edwards S.V., Pabo S., Villablanca F.X. and Wilson A.C. (1989):** Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proc. Natl. Acad. Sci. USA* **86**: 6196-6200.

**Koenemann S., Schram F.R. and Iliffe T.M. (2006):** Trunk segmentation patterns in Remipedia. *Crustaceana* **79**(5): 607-631.

**Koh H.L., Hallam T.G. and Lee H.L. (1997):** Combined effects of environmental and chemical stressors on a model *Daphnia* population. *Ecological Modelling* **103**: 19-32.

**Kondo R., Satta Y., Matsuura E.T., Ishiwa H., Takahata N. and Chigusa S.I. (1990):** Incomplete maternal transmission of mitochondrial DNA in *Drosophila*. *Genetics* **126**: 657-663.

**Konomi N., Lebwohl E. and Zhang D. (2002):** Comparison of DNA and RNA extraction methods for mummified tissues. *Molecular and Cellular Probes* **16**(6): 445-451.

**Kosel S., Grasbon-Frodl E.M., Arima K., Chimelli L., Hahn M., Hashizume Y., Hulette C., Ikeda K., Jacobsen P.F., Jones M., Kobayashi M., Love S., Mizutani T., Rosemberg S., Sasaki A., Smith T.W., Takahashi H., Vortmeyer A.O. and Graeber M.B. (2001):** Inter-laboratory comparison of DNA preservation in archival paraffin-embedded human brain tissue from participating centres on four continents. *Neurogenetics* **3**(3): 163-170.

**Koshiba M., Ogawa K., Hamazaki S., Sugiyama T., Ogawa O. and Kitajima T. (1993):** The effect of formalin fixation on DNA and the extraction of high-molecular-weight DNA from fixed and embedded tissues. *Pathol. Res. Practice* **189**(1): 66-72.

**Koslow J.A., Williams A. and Paxton J.R. (1997):** How many demersal fish species in the deep sea? A test of a method to extrapolate from local to global diversity. *Biodiversity and Conservation* **6**: 1523-1532.

**Koslow J.A., Boehlert G.W., Gordon J.D.M., Haedrich R.L., Lorance P. and Parin N. (2000):** Continental slope and deep-sea fisheries: implications for a fragile ecosystem. *ICES Journal of Marine Science* **57**(3): 548-557.

**Krajewski C., Driskell A.C., Baverstock P.R. and Braun M.J. (1992):** Phylogenetic relationships of the thylacine (Mammalia, Thylacinidae) among dasyuroid marsupials: evidence from cytochrome-b DNA sequences. *Proc. Roy. Soc. Lond. B* **250**: 19-27.

**Kraytsberg Y., Schwartz M., Brown T.A., Ebralidse K., Kunz W.S., Clayton D.A., Vissing J. and Khrapko K. (2004):** Recombination of human mitochondrial DNA. *Science* **304**: 981.

**Kreitman M. and Wayne M.L. (1994):** Organization of genetic variation at the molecular level: Lessons from *Drosophila*. pp. 157-183 *in* Schierwater B., Streit B., Wagner G.P. and DeSalle R., editors. ***Molecular ecology and evolution: Approaches and applications***. Birkhauser Verlag, Basel, Boston,

**Kress N., Hornung H. and Herut B. (1998):** Concentrations of Hg, Cd, Cu, Zn, Fe and Mn in deep sea benthic fauna from the Southeastern Mediterranean Sea: A comparison study between fauna collected at a pristine area and at two waste disposal sites. *Marine Pollution Bulletin* **36**(11): 911-921.

**Krings M., Stone A., Schmitz R.W., Krainitzki H., Stoneking M. and Paabo S. (1997):** Neandertal DNA sequences and the origin of modern humans. *Cell* **90**: 19-30.

**Kropp R.K. (2004):** *Review of deep-sea ecology and monitoring as they relate to deep-sea oil and gas operations*. Battelle Marine Sciences Laboratory, Washington, USA (Prepared for the U.S. Department of Energy).

**Krutovskii K.V., Vollmer S.S., Sorensen F.C., Adams W.T., Knapp S.J. and Strauss S.H. (1998):** RAPD genome maps of Douglas-Fir. *The Journal of Heredity* **89**(3): 197-205.

**Kullmann F., Scholmerich J., Bocker T. and Ruschoff J. (1998):** A comparison of methods for DNA extraction from paraffin-embadded tissue for microsatellite instability analysis by polymerase chain reaction. *Acta Biotechnologica* **18**(1): 77-83.

**Kumazawa Y. and Nishida M. (1995):** Variations in mitochendrial tRNA gene organization of reptiles as phylogenetic markers. *Mol. Biol. Evol.* **12**: 759-772.

**Kvist L., Martens J., Nazarenko A.A. and Orell M. (2003):** Paternal leakage of mitochondrial DNA in the great tit (*Parus major*). *Mol. Biol. Evol.* **20**: 243-247.

**Ladoukakis E.D. and Zouros E. (2001):** Recombination in animal mitochondrial DNA: Evidence from published sequences. *Mol. Biol. Evol.* **18**(11):2127–2131.

**Lan N., Montelione G.T. and Gerstein M. (2003):** Ontologies for proteomics: towards a systematic definition of structure and function that scales to the genome level. *Current Opinion in Chemical Biology* **7**: 44-54.

**Landsteiner K., Longsworth L.G. and Van der Scheer J. (1938):** Electrophoresis experiments with egg albumins and hemoglobins. *Science* **88**: 83-85.

**Lane M.A. (1996):** Roles of natural history collections. *Ann. Missouri Bot. Gard.* **83**: 536-545.

**Latham K.E. (2003):** Using amplification of bacteriophage lambda DNA to detect PCR inhibitors in skeletal DNA. *Electronic document on file at* http://archlab.uindy.edu, *University of Indianapolis, Archeology & Forensic Laboratory.* (This paper was originally presented at the 55th Meeting of the American Academy of Forensic Sciences, Chicago, Illinois, February 2003; *Proceedings of the American Academy of Forensic Sciences* **9**: 264)

**Lawrence E. (1995):** ***Dictionary of Biological Terms***. 11th edition. Addison Wesley Longman Limited, U.K.

**Lecointre G., Bonillo C., Ozouf-Costaz C. and Hureau J.-C. (1997):** Molecular evidence for the origins of Antarctic fishes: paraphyly of the Bovichtidae and no indication for the monophyly of the Notothenioidei (Teleostei). *Polar Biol.* **18**: 193-208.

**Lee C.I.P., Leong S.H., Png A.E.H., Choo K.W., Syn C., Lim D.T.H., Law H.Y. and Kon O.L. (2006):** An isothermal method for whole genome amplification of fresh and degraded DNA for comparative genomic hybridization, genotyping and mutation detection. *DNA Research* **13**: 77-88.

**Lee M.S.Y. (2004):** The molecularisation of taxonomy. *Invertebrate Systematics* **18**: 1-6.

**Legrand B., de Mazancourt P., Durigon M., Khalifat V. and Crainic K. (2002):** DNA genotyping of unbuffered formalin fixed paraffin embedded tissues. *Forensic Science International* **125**: 205-211.

**Levin L.A. (2006):** Recent progress in undrstanding larval dispersal: new directions and digressions. *Integrative and Comparative Biology* **46**(3): 282-297.

**Levin L.A., Etter R.J., Rex M.A., Gooday A.J., Smith C.R., Pineda J., Stuart C.T., Hessler R.R. and Pawson D. (2001):** Environmental influences on regional deep-sea species diversity. *Annu. Rev. Ecol. Syst.* **32:** 51-93.

**Li J., Liao X. and Yang H. (2000):** Molecular characterization of a parasitic tapeworm (*Ligula*) based on DNA sequences from formalin-fixed specimens. *Biochemical Genetics* **38**: 309-322.

**Li Y-C., Korol A.B., Fahima T., Beiles A. and Nevo E. (2002):** Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Molecular Ecology* **11**: 2453-2465.

**Li Y-C., Korol A.B., Fahima T. and Nevo E. (2004):** Microsatellites within genes: structure, function, and evolution. *Mol. Biol. Evol.* **21**(6): 991-1007.

**Liborio T.N., Etges A., Neves A.C., Mesquita R.A. and Nunes F.D. (2005):** Evaluation of the genomic DNA extracted from formalin-fixed, paraffin-embedded oral samples archived for the past 40-years. *J Bras Patol Med Lab* **41**(6): 405-410.

**Lilley M.D., Butterfield D.A., Olson E.J., Lupton J.E., Macko S.A. and McDuff R.E. (1993):** Anomalous $CH_4$ and $NH_4^+$ concentrations at an unsedimented mid-ocean-ridge hydrothermal system. *Nature, London* **364**: 45-47.

**Lindahl T. (1993a):** Recovery of antediluvian DNA. *Nature* **362**: 709-715.

**Lindahl T. (1993b):** Instability and decay of the primary structure of DNA. *Nature* **365**: 700.

**Linville J.G., Hayes J. and Wells J.D. (2004):** Mitochondrial DNA and STR analyses of maggot crop contents: effect of specimen preservation technique. *Journal of Forensic Sciences* **49**(2): 341-344.

**Lipscomb D., Platnick N. and Wheeler Q. (2003):** The intellectual content of taxonomy: a comment on DNA taxonomy. *Trends Ecol. Evol.* **18**: 65-66.

**Littlewood T. (1999):** Reply. *Parasitology Today* **15**(3): 122.

**Liu Y., Li C.M., Lu Z., Ding S., Yang X. and Mo J. (2006):** Studies on formation and repair of formaldehyde-damaged DNA by detection of DNA-protein crosslinks and DNA breaks. *Frontiers in Bioscience* **11**: 991-997.

**Liu Y.G., Liu L.X., Lei Z.W., Gao A.Y. and Li B.F. (2006):** Identification of polymorphic microsatellite markers from RAPD product in turbot (*Scophthalmus maximus*) and a test of cross-species amplification. *Molecular Ecology Notes* **6**: 867-869.

**Liu Z.J. and Cordes J.F. (2004):** DNA marker technologies and their applications in aquaculture genetics. *Aquaculture* **238**: 1-37.

**Long E.O. and Dawid I.B. (1980):** Repeated genes in Eukaryotes. *Annu. Rev. Biochem.* **49**: 727-764.

**Loreille O., Orlando L., Patou-Mathis M., Philippe M., Taberlet P. and Hanni C. (2001):** Ancient DNA analysis reveals divergence of the cave bear, *Ursus spelaeus*, and brown bear, *Ursus arctos*, lineages. *Current Biology* **11**: 200-203.

**Lougheed S.C., Gibbs H.L., Prior K.A. and Weatherhead P.J. (2000):** A comparison of RAPD versus microsatelite DNA markers in population studies of the massasauga rattlesnake. *The Journal of Heredity* **91**(6): 458-463.

**Lucchini V., Fabbri E., Marucco F., Ricci S., Boitani L. and Randi E. (2002):** Noninvasive molecular tracking of colonizing wolf (*Canis lupus*) packs in the western Italian Alps. *Molecular Ecology* **11**: 857-868.

**Luceri C., De Filippo C., Caderni G., Gambacciani L., Salvadori M., Giannini A. and Dolara P. (2000):** Detection of somatic DNA alterations in azoxymethane-induced F344 rat colon tumors by random amplified polymorphic DNA analysis. *Carcinogenesis* **21**(9): 1753-1756.

**Ludes B., Pfitzinger H. and Mangin P. (1993):** DNA fingerprinting from tissues after variable postmortem periods. *J. Forensic Sci* **38**(3): 686-690.

**Lundberg K.S., Shoemaker D.D., Adams M.W., Short J.M., Sorge J.A. and Mathur E.J. (1991):** High-fidelity amplification using a thermostable DNA polymerase isolated from *Pyrococcus furiosus*. *Gene* **108**: 1-6.

**Lunt D.H. and Hyman B.C. (1997):** Animal mitochondrial DNA recombination. *Nature* **387**: 247.

**Lunt D.H., Hutchinson W.F. and Carvalho G.R. (1999):** An efficient method for PCR-based isolation of microsatellite arrays (PIMA). *Molecular Ecology* **8**: 891-894.

**Luo L., Li B. and Pretlow T.P. (2003):** DNA alterations in human aberrant crypt foci and colon cancers by random primed polymerase chain reaction. *Cancer Research* **63**: 6166-6169.

**Lynch M., Koskella B. and Schaack S. (2006):** Mutation pressure and the evolution of organelle genomic architecture. *Science* **311**: 1727-1730.

**Macey J.R., Larson A., Ananjeva N.B., Fang Z. and Papenfuss T.J. (1997):** Two novel gene orders and the role of light-strand replication in rearrangement of the vertebrate mitochondrial genome. *Mol. Biol. Evol.* **14**(1): 91-104.

**Machugh D.E., Edwards C.J., Bailey J.F., Bancroft D.R. and Bradley D.G. (2000):** The extraction and analysis of ancient DNA from bone and teeth: a survey of current methodologies. *Ancient Biomolecules* **3**: 81-102.

**Mackie I.M., Pryde S., Gonzales-Sotelo C., Perez-Martin R., Quinteiro J. and Rey-Mendez M. (1995):** Fish species identification in canned tuna by DNA analysis (PCR-SSCP). *Inf. Fischwirtsch.* **42**(4): 207-210.

**Maclean N. (1987):** ***Macmillan Dictionary of Genetics and Cell Biology***. The Macmillan Press Ltd., U.K.

**Madurell T., Cartes J.E. and Labropoulou M. (2004):** Changes in the structure of fish assemblages in a bathyal site of the Ionian Sea (eastern Mediterranean). *Fisheries Research* **66**: 245-260.

**Maeda T., Jikko A., Hiranuma H. and Fuchihata H. (1999):** Analysis of genomic instability in squamous cell carcinome of the head and neck using the random amplified polymorphic DNA method. *Cancer Letters* **138**: 183-188.

**Magoulas A. and Zouros E. (1993):** Restriction-site heteroplasmy in anchovy (*Engraulis encrasicolus*) indicates incidental biparental inheritance of mitochondrial DNA. *Mol. Biol. Evol.* **10**: 319-325.

**Mahony J.B., Chong S., Coombes B.K., Smieja M. and Petrich A. (2000):** Analytical sensitivity, reproducibility of results, and clinical performance of five PCR assays for detecting *Chlamydia pneumoniae* DNA in peripheral blood mononuclear cells. *Journal of Clinical Microbiology* **38**(7): 2622-2627.

**Makowski G.S., Davis E.L. and Hopfer S.M. (1997):** Amplification of Guthrie card DNA: Effect of guanidine thiocyanate on binding of natural whole blood PCR inhibitors. *Journal of Clinical Laboratory Analysis* **11**: 87-93.

**Malhotra A. and Thrope R.S. (2004):** Maximizing information in systematic revisions: a combined molecular and morphological analysis of a cryptic green pitviper complex (*Trimeresurus stejnegeri*). *Biological Journal of the Linnean Society* **82**: 219-235.

**Mallet J. and Willmott K. (2003):** Taxonomy: renaissance or tower of Babel? *TRENDS in Ecology and Evolution* **18**(2): 57-59.

**Mandrioli M., Borsatti F. and Mola L. (2006):** Factors affecting DNA preservation from museum-collected lepidopteran specimens. *Entomologia Experimentalis et Applicata* **120**(3): 239-244.

**Markert C.L. and Moller F. (1959):** Multiple forms of enzymes: tissue, ontogenetic, and species specific patterns. *Proc. Natl. Acad. Sci. USA* **45**: 753-763.

**Markmann M. and Tautz D. (2005):** Reverse taxonomy: an approach towards determining the diversity of meiobenthic organisms based on ribosomal RNA signature sequences. *Philosophical Transactions: Biological Sciences* **360** (1462): 1917-1924.

**Marshall N.B. (1979):** ***Developments in Deep-Sea Biology***. Blandford Press, Poole.

**Marshall N.B. and Iwamoto T. (1973):** Family MACROURIDAE. *In:* ***Fishes of the Western North Atlantic.*** Number I, Part six. Edited by Cohen D.M. *et al.*, Memoir Sears Foundation for Marine Research, Yale University, New Haven.

**Martin E. and Hine R.S. (2000):** *A Dictionary of Biology*. Fourth Edition. Oxford University Press.

**Martin J.W. (2004):** Fixation versus preservation. *Web Page of Natural History Museum of Los Angeles County*. (http://clade.acnatsci.org/rosenberg/archiving/method/fixation_preservation.html).

**Maruyama T. and Birky C.W. Jr. (1991):** Effects of periodic selection on gene diversity in organelle genomes and other systems without recombination. *Genetics* **127**: 449-451.

**Mason B. (2003):** Marine survey sees net gain in number of fish species. *Nature* **425**: 889.

**Mason P.A. and Lightowlers R.N. (2003):** Hypothesis: Why do mammalian mitochondria possess a mismatch repair activity? *FEBS Letters* **554**: 6-9.

**Mayr E. (1996):** What is species, and what is not? *Philosophy of Science* **63**: 262-277.

**McCafferty S., Bermingham E., Quenouille B., Planes S., Hoelzer G. and Asoh K. (2002):** Historical biogeography and molecular systematics of the Indo-Pacific genus *Dascyllus* (Teleostei: Pomacentridae). *Molecular Ecology* **11**: 1377-1392.

**McClain C.R., Boyer A.G. and Rosenberg G. (2006):** The island rule and the evolution of body size in the deep sea. *Journal of Biogeography* **33**: 1578-1584.

**McDonald J.P., Hall A., Gasparutto D., Cadet J., Ballantyne J. and Woodgate R. (2006):** Novel thermostable Y-family polymerases: applications for the PCR amplification of damaged or ancient DNAs. *Nucleic Acids Research* **34**(4): 1102-1111.

**McMillan P.J. (1999):** New grenadier fishes of the genus *Coryphaenoides* (Pisces: Macrouridae); one from off New Zealand and one widespread in the southern Indo-West Pacific and Atlantic Oceans. *New Zealand Journal of Marine and Freshwater Research* **33**: 481-489.

**McNevin D., Wilson-Wilde L., Robertson J., Kyd J. and Lennard C. (2005):** Short tandem repeat (STR) genotyping of keratinised hair Part 2. An optimised genomic DNA extraction procedure reveals donor dependence of STR profiles. *Forensic Science International* **153**: 247-259.

**Mc Sherry E.A., Mc Goldrick A., Kay E.W., Hopkins A.M., Gallagher W.M. and Dervan P.A. (2007):** Formalin-fixed paraffin-embedded clinical tissues show spurious copy number changes in array-CGH profiles. *Clin Genet* **72**: 441-447.

**Mecocci P., MacGarvey U. and Beal M.F. (1994):** Oxidative damage to mitochondrial DNA is increased in Alzheimer`s disease. *Ann Neurol* **36**: 747-751.

**Medina M., Weil E. and Szmant A.M. (1999):** Examination of the *Montastraea annularis* species complex (Cnidaria: Scleractinia) using ITS and COI sequences. *Marine Biotechnology* **1**: 89-97.

**Mehling A., Wehmeier U.F. and Piepersberg W. (1995):** Application of random amplified polymorphic DNA (RAPD) assays in identifying conserved regions of actinomycete genomes. *FEMS Microbiology Letters* **128**: 119-126.

**Meland S., Johansen S., Johansen T., Haugli K. and Haugli F. (1991):** Rapid disappearance of one parental mitochondrial genotype after isogamous mating in the myxomycete *Physarum polycephalum*. *Curr. Genet.* **19**: 55-60.

**Mellado E., Diaz-Guerra T.M., Cuenca-Estrella M., Buendia V., Aspa J., Prieto E., Villagrasa J.R. and Rodriguez-Tudela J.-L. (2000):** Characterization of a possible nosocomial aspergillosis outbreak. *Clin Microbiol Infect* **6**: 543-548.

**Melton K., de Bakker M., Richardson M. and Witte F. (2006):** DNA extraction and amplification from formalin-preserved cichlids. *Prairie Universities Biological Symposium*, Calgary.

**Menzies R.J. (1965):** Condition for the existence of life on the abyssal sea floor. *Oceanography and Marine Biology Annual Review* **3**: 195-210.

**Menzies R.J., George R.Y. and Rowe G.T. (1973):** ***Abyssal environment and ecology of the World Oceans***. Wiley, New York.

**Merrett N.R. (1986):** Macrouridae of the eastern North Atlantic. *Fich. Ident. Plancton*, 173/174/175, 14 pp.

**Merrett N.R. (1989):** The elusive macrourid alevin and its seeming lack of potential in contributing to intrafamilial systematics. Pages 175-185 *in* Papers on the systematics of gadiform fishes. D.M. Cohen (ed.). *Los Ang. County Nat. Hist. Mus. Sci. Ser.* **32**: 262 pp

**Merrett N.R. (2003):** Preliminary guide to the identification of the early life history stages of bathygadid & macrourid fishes of the western central North Atlantic. *NOAA Technical Memorandum NMFS-SEFSC-498*, 32 p.

**Merrett N.R. and Marshall N.B. (1981):** Observation on the ecology of deep-sea bottom-living fishes collected off northwest Africa (08°-27°). *Prog. Oceanog.* **9**: 185-244.

**Merrett N.R. and Barnes S.H. (1996):** Preliminary survey of egg envelope morphology in the Macrouridae and the possible implications of its ornamentation. *Journal of Fish Biology* **48**: 101-119.

**Merrett N.R. and Haedrich R.L. (1997):** ***Deep-Sea demersal fish and fisheries***. Chapman & Hall.

**Merrett N.R. and Iwamoto T. (2000):** Pisces, Gadiformes: Grenadier fishes of the New Caledonian region, Southwest Pacific Ocean. Taxonomy and Distribution, with ecological notes. ***In:*** A. Crosnier (ed.), Resultats des Campagnes MUSORSTOM, Volume 21. *Memoires du Museum nationale d'Histoire naturelle*, **184:** 723-781.

**Merrett N.R. and Moore J.A. (2005):** A new genus and species of deep demersal fish (Teleostei: Stephanoberycidae) from the tropical eastern North Atlantic. *Journal of Fish Biology* **67**: 1699-1710.

**Meyer A. (1993):** Chapter 1: Evolution of mitochondrial DNA in fishes. pp. 1-38 *in* Hochachka P.W. and Mommsen T.P., editors. ***Biochemistry and molecular biology of fishes.*** Elsevier Science Publishers, Amsterdam

**Micheli M.R., Bova R., Pascale E. and D`Ambrosio E. (1994):** Reproducible DNA fingerprinting with the random amplified polymorphic DNA (RAPD) method. *Nucleic Acids Research* **22**(10): 1921-1922.

**Miething F., Hering S., Hanschke B. and Dressler J. (2006):** Effect of fixation to the degradation of nuclear and mitochondrial DNA in different tissues. *Journal of Histochemistry and Cytochemistry* **54**(3): 371-374.

**Mindell D. (1991):** Aligning DNA sequences: Homology and phylogenetic weighting. *In:* M. Miyamoto and J. Cracraft, eds. ***Phylogenetic analysis of DNA sequences.*** Oxford University Press, Oxford. pp. 73-89

**Mitchell D., Willerslev E. and Hansen A. (2005):** Damage and repair of ancient DNA. *Mutation Research* **571**: 265-276.

**Miya M. and Nishida M. (1996):** Molecular phylogenetic perspective on the evolution of the deep-sea fish genus *Cyclothone* (Stomiiformes: Gonostomatidae). *Ichthyological Research* **43**: 375-398.

**Miya M. and Nishida M. (1999):** Organization of the mitochondrial genome of a deep-sea fish, *Genostoma gracile* (Teleostei: Stomiiformes): First example of transfer RNA gene rearrangements in bony fishes. *Mar Biotechnol* **1**(5): 416-426.

**Miya M., Takeshima H., Endo H., Ishiguro N.B., Inoue J.G., Mukai T., Satoh T.P., Yamaguchi M., Kawaguchi A., Mabuchi K., Shirai M. and Nishida M. (2003):** Major patterns of higher teleostean phylogenies: a new perspective based on 100 complete mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution* **26**: 121-138.

**Miyata M., Inglis V. and Aoki T. (1996):** Rapid identification of *Aeromonas salmonicida* subspecies *salmonicida* by the polymerase chain reaction. *Aquaculture* **141**: 13-24.

**Molin D.G.M. and van Dam A.J. (2006):** A preservation method supporting multipurpose analysis of long-stored samples. *Cell Preservation Technology* **4**(1): 46-50.

**Monaghan M.T., Balke M., Gregory T.R. and Vogler A.P. (2005):** DNA-based species delineation in tropical beetles using mitochondrial and nuclear markers. *Philosophical Transactions: Biological Sciences* **360** (1462): 1925-1933.

**Monteiro P., Araujo A., Erzini K. and Castro M. (2001):** Discards of the Algarve (southern Portugal) crustacean trawl fishery. *Hydrobiologia* **449**: 267-277.

**Montiel R., Malgosa A. and Subira E. (1997):** Overcoming PCR inhibitors in ancient DNA extracts from teeth. *Ancient Biomolecules* **1**(3): 221-225.

**Moore J.L., Aros M., Steudel K.G. and Cheng K.C. (2002):** Fixation and decalcification of adult zebrafish for histological, immunocytochemical, and genotypic analysis. *BioTechniques* **32**(2): 296-298.

**Moran P., Kornfield I. and Reinthal P.N. (1994):** Molecular systematics and radiation of the haplochromine cichlids (Teleostei: Perciformes) of Lake Malawi. *Copeia* **1994**: 274-288.

**Moran P. and Baker J. (2002):** Polymerase chain reaction inhibition in historical fish scale mounting cards. *Transactions of the American Fisheries Society* **131**: 109-119.

**Morin P.A., Chambers K.E., Boesch C. and Vigilant L. (2001):** Quantitative polymerase chain reaction analysis of DNA from noninvasive samples for accurate microsatellite genotyping of wild chimpanzees (*Pan troglodytes verus*). *Molecular Ecology* **10**: 1835-1844.

**Morin P.A., LeDuc R.G., Robertson K.M., Hedrick N.M., Perrin W.F., Etnier M., Wade P. and Taylor B.L. (2006):** Genetic analysis of killer whale (*Orcinus orca*) historical bone and tooth samples to identify western U.S. ecotypes. *Marine Mammal Science* **22**(4): 897-909.

**Morin P.A. and Mc Carthy M. (2007):** Highly accurate SNP genotyping from histrorical and low-quality samples. *Marine Ecology Notes* **7**: 937-946.

**Morita T. (1999):** Molecular phylogenetic relationships of the deep-sea fish genus *Coryphaenoides* (Gadiformes: Macrouridae) based on mitochondrial DNA. *Molecular Phylogenetics and evolution* **13**(3): 447-454.

**Morita T. (2003):** Structure-based analysis of high pressure adaptation of α-actin. *The Journal of Biological Chemistry* **278**(30): 28060-28066.

**Moritz C. and Brown W.M. (1987):** Tandem duplications in animal mitochondrial DNAs: Variation in incidence and gene content among lizards. *Proc. Natl. Acad. Sci. USA* **84**: 7183-7187.

**Moritz C., Dowling T.E. and Brown W.M. (1987):** Evolution of animal mitochondrial DNA: relevance for population biology and systematics. *Ann. Rev. Ecol. Syst.* **18**: 269-292.

**Moritz C. and Cicero C. (2004):** DNA barcoding: Promise and pitfalls. *PLoS Biology* **2** (10): 1529-1531.

**Morley S.A., Mulvey T., Dickson J. and Belchier M. (2004):** The biology of the bigeye grenaider at South Georgia. *Journal of Fish Biology* **64**: 1514-1529.

**Mormede S. and Davis I.M. (2001):** Trace elements in deep-water fish species from the Rockall Trough. *Fisheries Research* **51**: 197-206.

**Mortensen P.B., Hovland M., Brattegard T. and Farestveit R. (1995):** Deep-water bioherms of the scleractinian coral *Lophelia pertusa* (L.) at 64° on the Norwegian shelf: structure and associated megafauna. *Sarsia* **80**: 145-158.

**Moura C.J., Harris D.J., Cunha M.R. and Rogers A.D. (2008):** DNA barcoding reveals cryptic diversity in marine hydroids (Cnidaria, Hydrozoa) from coastal and deep-sea environments. *Zoologica Scripta* **37**: 93-108.

**Mourier T. (2005):** Reverse transcription in genome evolution. *Cytogenetic and Genome Research* **110**: 56-62.

**Mtambo J., van Bortel W., Madder M., Roelants P. and Backeljau T. (2006):** Comparison of preservation methods of *Rhipicephalus appendiculatus* (Acari: Ixodidae) for reliable DNA amplification by PCR. *Experimental and Applied Acarology* **38**: 189-199.

**Mueller R.L. and Boore J.L. (2005):** Molecular mechanisms of extensive mitochondrial gene rearrangement in plethodontid salamanders. *Mol. Biol. Evol.* **22**(10): 2104-2112.

**Muller W.A. (1994):** To what extent does genetic information determine structural characteristics and document homologies? pp. 551-558 *in* Schierwater B., Streit B., Wagner G.P. and DeSalle R., editors. ***Molecular ecology and evolution: Approaches and applications***. Birkhauser Verlag, Basel, Boston, Berlin.

**Mulligan C.J. (2005):** Isolation and analysis of DNA from archaelogical, clinical, and natural history specimens. *Methodes in Enzymology* **395**: 87-103.

**Mullis K.B. and Faloona F.A. (1987):** Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. *Methods Enzymol* **155**: 335-350.

**Munasinghe D.H.N., Murphy N.P. and Austin C.M. (2003):** Utility of mitochondrial DNA sequences from four gene regions for systematic studies of Australian freshwater crayfish of the genus *Cherax* (Decapoda: Parastacidae). *Journal of Crustacean Biology* **23**(2): 402-417.

**Museum Victoria Australia (2003):** What is ancient DNA? (*http://www.museum.vic.gov.au/scidiscovery/dna/ancient_dna.asp*)

**Nadler S. (1999):** Nucleotide sequences from vintage helminths: Fine wine or vinegar? *Parasitology Today* **15**(3): 122.

**Nakanishi M., Wilson A.C., Nolan R.A., Gorman G.C. and Bailey G.S. (1969):** Phenoxyethanol: Protein preservative for taxonomists. *Science* **163**: 681-683.

**Needleman S.B. and Wunsch C.D. (1970):** A general method applicable to the search for similarities in the amino acid sequences of two proteins. *J. Mol. Biol.* **48**: 443-453.

**Nei M. and Tajima F. (1981):** DNA polymorphism detectable by restriction endonucleases. *Genetics* **105**: 207-217.

**Nesbo C.L., Arab M.O. and Jakobsen K.S. (1998):** Heteroplasmy, length and sequence variation in the mtDNA control regions of three percid fish species (*Perca fluviatilis, Acerina cernua, Stizostedion lucioperca*). *Genetics* **148**(4): 1907-1919.

**Newmeyer D.D. and Ferguson-Miller S. (2003):** Mitochondria: Releasing power for life and unleashing the machineries of death. *Cell* **112**: 481-490.

**Nitta Y., Tanaka H., Masuda Y. and Hoshi M. (2002):** The quality of DNA recovered from the archival tissues of atomic bomb survivors is good enough for the single nucleotide polymorphism analysis in spite of the decade-long preservation in formalin. *J. Radiat. Res.* **43**: 65-72.

**Nunes F. and Norris R.D. (2006):** Abrupt reversal in ocean overturning during the Palaeocene/Eocene warm period. *Nature* **439**: 60-63.

**Nyakaana S., Arctander P. and Siegismund H.R. (2002):** Population structure of the African savannah elephant inferred from mitochondrial control region sequences and nuclear microsatellite loci. *Heredity* **89**: 90-98.

**Nybom H. (2004):** Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology* **13**: 1143-1155.

**Olabarria C. (2005):** Patterns of bathymetric zonation of bivalves in the Porcupine Seabight and adjacent Abyssal plain, NE Atlantic. *Deep-Sea Research I* **52**: 15-31.

**O'Leary J.J., Browne G., Landers R.J., Crowley M., Healy I.B., Street J.T., Pollock A.M., Murphy J., Johnson M.I., Lewis F.A., Mohamdee O., Cullinane C. and Doyle C.T. (1994):** The importance of fixation procedures on DNA template and its suitability for solution-phase polymerase chain reaction and PCR *in situ* hybridization. *The Histochemical Journal* **26**(4): 337-346.

**Oliveira C.M., Fungaro M.H.P., Camargo L.E.A. and Lopes J.R.S. (2002):** Comparative analysis of the stability of DNA of *Dalbulus maidis* (DeLong & Wolcott) (Hemiptera: Cicadellidae) under different methods of preservation for use in RAPD-PCR. *Neotropical Entomology* **31**: 225-231.

**Olson L.E., Sargis E.J. and Martin R.D. (2005):** Intraordinal phylogenetics of treeshrews (Mammalia: Scandentia) based on evidence from the mitochondrial 12S rRNA gene. *Molecular Phylogenetics and Evolution* **35**(3): 656-673.

**Ong T., Song B., Qian H., Wu Z.L. and Whong W. (1998):** Detection of genomic instability in lung cancer tissues by random amplified polymorphic DNA analysis. *Carcinogenesis* **19**(1): 233-235.

**Orlando V., Strutt H. and Paro R. (1997):** Analysis of chromatin structure by *in vivo* formaldehyde cross-linking. *Methods* **11**(2): 205-214.

**O'Rourke D.H., Hayes M.G. and Carlyle S.W. (2000):** Ancient DNA studies in physical anthropology. *Annual Review of Anthropology* **29**: 217-242.

**Paabo S. (1993):** Ancient DNA. *Sci. Amer.* **269**: 60-66.

**Paabo S., Gifford J.A. and Wilson A.C. (1988):** Mitochondrial DNA sequences from 7000-year old brain. *Nucl. Acid Res.* **16**: 9775-9787.

**Paabo S., Irwin D.M. and Wilson A.C. (1990):** DNA damage promotes jumping between templates during enzymatic amplification. *The Journal of Biological Chemistry* **265**(8): 4718-4721.

**Paabo S., Poinar H., Serre D., Jaenicke-Despres V., Hebler J., Rohland N., Kuch M., Krause J., Vigilant L. and Hofreiter M. (2004):** Genetic analyses from ancient DNA. *Annu. Rev. Genet.* **38**: 645-679.

**Page T.J., Choy S.C. and Hughes J.M. (2005):** The taxonomic feedback loop: symbiosis of morphology and molecules. *Biology Letters* **1**: 139-142.

**Palumbi S., Martin A., Romano S., McMillan W.O., Stice L. and Grabowski G. (1991):** ***The Simple Fool`s Guide to PCR***. University of Hawaii, Honolulu, HI.

**Palumbi S.R. (1994):** Genetic divergence, reproductive isolation, and marine speciation. *Annu. Rev. Ecol. Syst.* **25**: 547-572.

**Palumbi S.R. (1996a):** Nucleic acids II: The polymerase chain reaction. pp. 205-247 *in* Hillis D.M., Moritz C. and Mable B.K. (eds.), ***Molecular Systematics.*** Second edition. Sinauer Associates, Inc., Publishers Sunderland, Massachusetts, U.S.A.

**Palumbi S.R. (1996b):** What can molecular genetics contribute to marine biogeography? An urchin's tale. *Journal of Experimental Marine Biology and Ecology* **203**(1): 75-92.

**Pandey M., Gailing O., Fischer D., Hattemer H.H. and Finkeldey R. (2004):** Characterization of microsatellite markers in sycamore (*Acer pseudoplatanus* L.). *Molecular Ecology Notes* **4**(2): 253-255.

**Paradis Y., Brodeur P., Mingelbier M. and Magnan P. (2007):** Length and weight reduction in larval and juvenile yellow perch preserved with dry ice, formalin, and ethanol. *North American Journal of Fisheries Management* **27**(3): 1004-1009.

**Paran I. and Michelmore R.W. (1993):** Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. *Theor. Appl Genet.* **85**: 985-993.

**Parham J.F., Stuart B.L., Bour R. and Fritz U. (2004):** Evolutionary distinctiveness of the extinct Yunnan box turtle (*Cuora yunnanensis*). *Proc Biol Sci.* **271**(Suppl 6): S391-S394.

**Park J.W., Beaty T.H., Boyce P., Scott A.F. and McIntosh I. (2005):** Comparing whole-genome amplification methods and sources of biological samples for single-nucleotide polymorphism genotyping. *Clinical Chemistry* **51**(8): 1520-1523.

**Park L.K. and Moran P. (1995):** Developments in molecular genetic techniques in fisheries. pp. 1-28 *in:* G.R. Carvalho and T.J. Pitcher, editors. ***Molecular genetics in fisheries***. First edition. Chapman & Hall.

**Patel P.H., Kawate H., Adman E., Ashbach M. and Loeb L.A. (2001):** A single highly mutable catalytic site amino acid is critical for DNA polymerase fidelity. *J Biol Chem* **276**: 5044-5051.

**Patton T.G., Katz S., Sobieski R.J. and Crupper S.S. (2001):** Genotyping of clinical *Serratia marcescens* isolates: a comparison of PCR-based methods. *FEMS Microbiology Letters* **194**: 19-25.

**Pavelic J., Gall-Troselj K., Bosnar M.H., Kardum M.M. and Pavelic K. (1996):** PCR amplification of DNA from archival specimens. A methodological approach. *Neoplasma* **43**(2): 75-81.

**Payne R.B. and Sorenson M.D. (2002):** Museum collections as sources of genetic data. *Bonner Zoologische Beitrage* **51**: 97-104.

**Pearson C.V.M., Rogers A.D. and Sheader M. (2002):** The genetic structure of the rare lagoonal sea anemone, *Nematostella vectensis* (Cnidaria; Anthozoa) in the United Kingdom based on RAPD analysis. *Molecular Ecology* **11**: 2285-2293.

**Peck L.S., Clark M.S., Clarke A., Cockell C.S., Convey P., Detrich III H.W., Fraser K.P.P., Johnston I.A., Methe B.A., Murray A.E., Romisch K. and Rogers A.D. (2005):** Genomics: applications to Antarctic ecosystems. *Polar Biol* **28**: 351-365.

**Peek A.S., Gaut B.S., Feldman R.A., Barry J.P., Kochevar R.E., Lutz R.A., Vrijenhock R.C. (2000):** Neutral and nonneutral mitochondrial genetic variation in deep-sea clams from the family Vesicomydae. *J. Mol. Evol.* **50**: 141-153.

**Peijnenburg K.T.C.A., van Haastrecht E.K. and Fauvelot C. (2005):** Present-day genetic composition suggests contrasting demographic histories of two dominant chaetognaths of the North-East Atlantic, *Sagitta elegans* and *S. setosa*. *Marine Biology* **147**: 1279-1289.

**Pennec J.-P, Wardle C.S., Harper A.A. and Macdonald A.G. (1988):** Effects of high hydrostatic pressure on the isolated hearts of shallow water and deep sea fish: results of *Challenger* cruise 6B/85. *Comparative Biochemistry and Physiology* **89A,2**: 215-218.

**Penner G.A., Bush A., Wise R., Kim W., Domier L., Kasha K., Laroche A., Scoles G., Molnar S.J. and Fedak G. (1993):** Reproducibility of random amplified polymorphic DNA (RAPD) analysis among laboratories. Cold Spring Harbor Press, NY. *PCR Methods and Applications* **2**: 341-345.

**Perez J., Alvarez P., Martinez J.L. and Garcia-Vazquez E. (2005):** Genetic identification of hake and megrim eggs in formaldehyde-fixed plankton samples. *ICES Journal of Marine Science* **62**: 908-914.

**Pergams O.R.W., Barnes W.M. and Nyberg D. (2003):** Rapid change in mouse mitochondrial DNA. *Nature* **423**: 397-398.

**Perry W.L. 3rd, Bass W.M. and Riggsby W.S. (1988):** The autodegradation of deoxyribonucleic acid (DNA) in human rib bone and its relationship to the time interval since death. *J. Forensic Sci.* **33**(1): 144-153.

**Pettersson E., Lindskog M., Lundeberg J. and Ahmadian A. (2006):** Tri-nucleotide threading for parallel amplification of minute amounts of genomic DNA. *Nucleic Acids Research* **34**(6): e49.

**Pettitt C. (1994):** Using the collections. Chapter 6 *in* Stansfield G., Mathias J. and Reid G., editors. ***Manual of natural history curatorship.*** HMSO, London, UK.

**Piganeau G., Gardner M. and Eyre-Walker A. (2004):** A broad survey of recombination in animal mitochondria. *Mol. Biol. Evol.* **21**(12): 2319-2325.

**Piggott M.P. and Taylor A.C. (2003):** Remote collection of animal DNA and its applications in conservation management and understanding the population biology of rare and cryptic species. *Wildlife Research* **30**(1): 1-13.

**Poinar H.N. (2002):** The genetic secrets some fossils hold. *Acc. Chem. Res.* **35**: 676-684.

**Poinar H.N. (2003):** The top 10 list: criteria of authenticity for DNA from ancient and forensic samples. *International Congress Series* **1239**: 575-579.

**Poinar H.N., Schwarz C., Qi J., Shapiro B., MacPhee R.D.E., Buigues B., Tikhonov A., Huson D. H., Tomsho L.P., Auch A., Rampp M., Miller W. and Schuster S.C. (2006):** Metagenomics to paleogenomics: Large-Scale sequencing of mammoth DNA. *Science* **311**: 392-394.

**Poljak M., Seme K. and Gale N. (2000):** Rapid extraction of DNA from archival clinical specimens: our experiences. *Pflugers Archiv European Journal of Physiology* **439**(7): 1432-2013.

**Porceddu A., Albertini E., Barcaccia G., Marconi G., Bertoli F.B. and Veronesi F. (2002):** Development of S-SAP markers based on an LTR-like sequence from *Medicago sativa* L. *Mol Genet Genomics* **267**: 107-114.

**Post R.J., Flook P.K. and Millest A.L. (1993):** Methods for the preservation of insects for DNA studies. *Biochem. Syst. Ecol.* **21**: 85-92.

**Potts W.K. (1996):** PCR-based cloning across large taxonomic distances and polymorphism detection: MHC as a case study. pp. 181-194 *in: Molecular Zoology: Advances, Strategies, and Protocols.* J.D. Ferraris and S.R. Palumbi, editors. Wiley-Liss, New York.

**Prak E.T.L. and Kazazian H.H., Jr. (2000):** Mobile elements and the human genome. *Nature Reviews/ Genetics* **1**: 134-144.

**Prendini L., Hanner R. and DeSalle (2002):** Obtaining, storing and archiving specimens and tissue samples for use in molecular studies (Chapter 11). *In:* DeSalle R., Giribet G. and Wheeler W., editors. *Techniques in molecular systematics and evolution*. Birkhauser Verlag, Basel, Boston, Berlin.

**Prioli S.M.A.P., Prioli A.J., Julio H.F., Jr., Pavanelli C.S., de Oliveira A.V., Carrer H., Carraro D.M. and Prioli L.M. (2002):** Identification of *Astyanax altiparanae* (Teleostei, Characidae) in the Iguacu River, Brasil, based on mitochondrial DNA and RAPD markers. *Genetics and Molecular Biology* **25**(4): 421-430.

**Probert P.K. (1999):** Seamounts, sanctuaries and sustainability: moving towards deepsea conservation. *Aquatic conservation: Marine and Freshwater Ecosystems* **9**: 601-605.

**Probert P.K., McKnight D.G. and Grove S.L. (1997):** Benthic invertebrate bycatch from a deepwater trawl fishery, Chathm Rise, New Zealand. *Aquatic conservation: Marine and Freshwater Ecosystems* 7: 27-40.

**Proudlove G. and Wood P.J. (2003):** The blind leading the blind: cryptic subterranean species and DNA taxonomy. *Trends in Ecology and Evolution* **18**(6): 272-273.

**Pruvost M., Grange T. and Geigl E-M. (2005):** Minimizing DNA contamination by using UNG-coupled quantitative real-time PCR on degraded DNA samples: application to ancient DNA studies. *BioTechniques* **38**(4): 569-575.

**Pusch C.M. and Bachmann L. (2004):** Spiking of contemporary human template DNA with ancient DNA extracts induces mutations under PCR and generates nonauthentic mitochondrial sequences. *Molecular Biology and Evolution* **21**(5): 957-964.

**QIAGEN (2002):** 8. Effect of template quality on PCR performance. *In: Critical Factors for Successful PCR*. Qiagen.

**Quach N., Goodman M.F. and Shibata D. (2004):** *In vitro* mutation artifacts after formalin fixation and error prone tanslesion synthesis during PCR. *BMC Clinical Pathology* 4:1 (http://www.biomedcentral.com/1472-6890/4/1)

**Quattro J.M., Chase M.R., Rex M.A., Greig T.W. and Etter R.J. (2001):** Extreme mitochondrial DNA divergence within populations of the deep-sea gastropod *Frigidoalvania brychia*. *Marine Biology* **139**(6): 1107-1113.

**Quievryn G. and Zhitkovich A. (2000):** Loss of DNA-protein crosslinks from formaldehyde-exposed cells occurs through spontaneous hydrolysis and an active repair process linked to proteosome function. *Carcinogenesis* **21**(8): 1573-1580.

**Rabouam C., Comes A.M., Bretagnolle V., Humbert J-F., Periquets G. and Bigots Y. (1999):** Features of DNA fragments obtained by random amplified polymorphic DNA (RAPD) assays. *Molecular Ecology* **8**: 493-503.

**Radstrom P., Knutsson R., Wolffs P., Lovenklev M. and Lofstrom C. (2004):** Pre-PCR processing – Strategies to generate PCR-compatible samples. *Molecular Biotechnology* **26**(2): 133-146.

**Rainbow P. (2001):** Collections: past, present and future. In: *Marine sample collections: their value, use and future.* (Prepared for IACMST by Rothwell R.G.). *IACMST Information Document No. 8*, July 2001.

**Rait V.K., O`Leary T.J. and Mason J.T. (2004):** Modeling formalin fixation and antigen retrival with bovine pancreatic ribonuclease A: I-Structural and functional alterations. *Laboratory Investigation* **84**: 292-299.

**Rait V.K., Zhang Q., Fabris D., Mason J.T. and O`Leary T.J. (2006):** Conversions of formaldehyde-modified 2`-deoxyadenosine 5`-monophosphate in conditions modeling formalin-fixed tissue dehydration. *Journal of Histochemistry and Cytochemistry* **54**(3): 301-310.

**Rajput S.G., Wable K.J., Sharma K.M., Kubde P.D. and Mulay S.A. (2006):** Reproducibility testing of RAPD and SSR markers in tomato. *African Journal of Biotechnology* **5**(2): 108-112.

**Ramser J., Weising K., Chikaleke V. and Kahl G. (1997):** Increased informativeness of RAPD analysis by detection of microsatellite motifs. *BioTechniques* **23**: 285-290.

**Ran Y., Hammett K.R.W. and Murray B.G. (2001):** Phylogenetic analysis and karyotype evolution in the genus *Clivia* (Amaryllidaceae). *Annals of Botany* **87**: 823-830.

**Rand D.M. (1993):** Endotherms, ectotherms and mitochondrial genome-size variation. *J. Mol. Evol.* **37**: 281-295.

**Rand D.M. (1994):** Concerted evolution and RAPing in mitochondrial VNTRs and the molecular geography of cricket populations. pp. 227-245 *in* Schierwater B., Streit B., Wagner G.P. and DeSalle R., editors. ***Molecular ecology and evolution: Approaches and applications***. Birkhauser Verlag, Basel, Boston, Berlin.

**Rand D.M. and Harrison R.G. (1989):** Molecular population genetics of mtDNA size variation in crickets. *Genetics* **121**: 551-569.

**Raupach M.J. and Wagele J-W. (2006):** Distinguishing cryptic species in Antarctic Asellota (Crustacea: Isopoda) – a preliminary study of mitochondrial DNA in *Acanthaspidia drygalskii*. *Antarctic Science* **18**(2): 191-198.

**Reeb A.C. and Avise J.C. (1990):** A genetic discontinuity in a continuously distributed species: mitochondrial DNA in the American oyster, Crassostrea virginica. *Genetics* **124**: 397-406.

**Reiss R.A., Schwert D.P. and Ashworth A.C. (1995):** Field preservation of Coleoptera for molecular-genetic analyses. *Environmental Entomology* **24**(3): 716-719.

**Reiss R.A. and Rutz B. (1999):** Quality control PCR: A method for detecting inhibitors of *Taq* DNA polymerase. *BioTechniques* **27**: 920-926.

**Remigio E.A. and Hebert P.D.N. (2003):** Testing the utility of partial COI sequence for phylogenetic estimates of gastropod relationships. *Molecular Phylogenetics and Evolution* **29**: 641-647.

**Reusch T.B.H., Stam W.T. and Olsen J.L. (2000):** A microsatellite-based estimation of clonal diversity and population subdivision in *Zostera marina*, a marine flowering plant. *Molecular Ecology* **9**(2): 127-140.

**Rex M.A. (1997):** An oblique slant on deep-sea biodiversity. *Nature* **385**: 577-578.

**Rex M.A., McClain C.R., Johnson N.A., Etter R.J., Allen J.A., Bouchet P. and Waren A. (2005a):** A source-sink hypothesis for abyssal biodiversity. *The American Naturalist* **165**(2): 163-178.

**Rex M.A., Crame J.A., Stuart C.T. and Clarke A. (2005b):** Large-scale biogeographic patterns in marine mollusks: A confluence of history and productivity? *Ecology* **86**(9): 2288-2297.

**Rey I., Dorda B.A. and Valdecasas A.G. (2004):** Traditional water mite fixatives and their compatibility with later DNA studies. *Experimental and Applied Acarology* **34**: 59-65.

**Ribeiro G.R.H., Francisco G., Teixeira L.V.S., Romao-Correia R.F., Sanches J.A. Jr., Neto C.F. and Ruiz I.R.G. (2004):** Repetitive DNA alterations in human skin cancers. *Journal of Dermatological Science* **36**: 79-86.

**Rieseberg L.H. (1996):** Homology among RAPD fragments in interspecific comparisons. *Molecular Ecology* **5**: 99-105.

**Rigaa A., Monnerot M. and Sellos D. (1995):** Molecular cloning and complete nucleotide sequence of the repeated unit and flanking gene of the scallop *Pecten maximus* mitochondrial DNA: Putative replication origin features. *Journal of Molecular Evolution* **41**(2): 189-195.

**Ristaino J.B. (1998):** The importance of archival and herbarium materials in understanding the role of oospores in late blight epidemics of the past. *Phytopathology* **88**(11): 1120-1130.

**Ristaino J.B., Parra G.R. and Groves C.T. (2000):** PCR amplification of *Phytophthora infestans* from 19th century herbarium specimens. *GILB Newsletter* No.10.

**Ritland K. (2005):** Multilocus estimation of pairwise relatedness with dominant markers. *Molecular Ecology* **14**: 3157-3165.

**Rivers P.J. and Ardren W.R. (1998):** The value of archives. *Fisheries* **23**: 6-9.

**Rocco L., Costagliola D., Fiorillo M., Tinti F. and Stingo V. (2005):** Molecular and chromosomal analysis of ribosomal cistrons in two cartilaginous fish, *Taeniura lymma* and *Raja montagui* (Chondrichthyes, Batoidea). *Genetica* **123**: 245-253.

**Rodriguez-Grana L., Herrera G., Herrera L. and Castro L.R. (2004):** Divergence of two forms of *Triphoturus* in the easteren Pacific based on mtDNA cytochrome b gene sequence and larval morphology. *Journal of Fish Biology* **64**(5): 1455-1461.

**Roeder A.D., Archer F.I., Poinar H.N. and Morin P.A. (2004):** A novel method for collection and preservation of faeces for genetic studies. *Molecular Ecology Notes* **4**: 761-764.

**Rogers A.D. (1999):** The biology of *Lophelia pertusa* (Linnaeus 1758) and other deep-water reef-forming corals and impacts from human activities. *Int. Rev. Hydrobiol.* **84**: 315-406.

**Rogers A.D. (2000):** The role of the oceanic oxygen minima in generating biodiversity in the deep sea. *Deep-Sea Research II* **47**: 119-148.

**Rogers A.D. (2001):** Molecular methods for marine biodiversity research and the use of preserved materials from museum collections. In: *Marine sample collections: their value, use and future.* (Prepared for IACMST by Rothwell R.G.). *IACMST Information Document No. 8*, July 2001.

**Rogers A.D., Morley S., Fitzcharles E., Jarvis K. and Belchier M. (2006):** Genetic structure of Patagonian toothfish (*Dissostichus eleginoides*) populations on the Patagonian Shelf and Atlantic and western Indian Ocean Sectors of the Southern Ocean. *Marine Biology* **149**(4): 915-924.

**Rogers S.O., Langenegger K. and Holdenrieder O. (2000):** DNA changes in tissues entrapped in plant resins (the precursors of amber). *Naturwissenschaften* **87**: 70-75.

**Rohland N., Siedel H. and Hofreiter M. (2004):** Nondestructive DNA extraction method for mitochondrial DNA analyses of museum specimens. *BioTechniques* **36**: 814-821.

**Rokas A., Ladoukakis E. and Zouros E. (2003):** Animal mitochondrial DNA recombination revisited. *J Biol Chem* **271**: 27536-27543.

**Romero R.L., Juston A.C., Ballantyne J. and Henry B.E. (1997):** The applicability of formalin-fixed and formalin fixed paraffin embedded tissues in forensic DNA analysis. *J. Forensic Sci* **42** (4): 708-714.

**Roques S., Fox C.J., Villasana M.I. and Rico C. (2006):** The complete mitochondrial genome of the whiting, *Merlangius merlangus* and the haddock, *Melanogrammus aeglefinus*: A detailed genomic comparison among closely related species of the Gadidae family. *Gene* **383**: 12-23.

**Ross K.G., Shoemaker D.D., Krieger M.J.B., DeHeer C.J. and Keller L. (1999):** Assessing genetic structure with multiple classes of molecular markers: A case study involving the introduced fire ant *Solenopsis invicta*. *Mol. Biol. Evol.* **16**(4): 525-543.

**Ross S.W., Sulak K.J. and Munroe T.A. (2001):** Association of *Syscenus infelix* (Crustacea: Isopoda: Aegidae) with benthopelagic rattail fishes, *Nezumia* spp. (Macrouridae), along the western North Atlantic continental slope. *Marine Biology* **138**: 595-601.

**Rossen L., Norskov P., Holmstrom K. and Rasmussen O.F. (1992):** Inhibition of PCR by components of food samples, microbial diagnostic assays and DNA-extraction solutions. *International Journal of Food Microbiology* **17**(1): 37-45.

**Routman E. and Cheverud J.M. (1994):** Individual genes underlying quantitative traits: Molecular and analytical methods. pp. 593-606 *In:* Schierwater B., Streit B., Wagner G.P. and DeSalle R., editors. ***Molecular ecology and evolution: Approaches and applications***. Birkhauser Verlag, Basel, Boston, Berlin.

**Roy M.S., Girman D.J., Taylor A.C. and Wayne R.K. (1994):** The use of museum specimens to reconstruct the genetic variability and relationships of extinct populations. *Experientia* **50**: 551-557.

**Rudin N. and Inman K. (2002):** *An introduction to forensic DNA analysis*. (2nd edition). CRC Press, Boca Raton, London, New York, Washington.

**Rundle H.D. and Nosil P. (2005):** Ecological speciation. *Ecology Letters* **8**: 336-352.

**Saiki R.K., Scharf S., Faloona F., Mullis K.B., Horn G.T., Erlich H.A. and Arnheim N. (1985):** Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**: 1350-1354.

**Saiki R.K., Gelfand D.H., Stoffel S., Scharf S.J., Higuchi R., Horn G.T., Mullis K.B., Erlich H.A. (1988):** Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**: 487-491.

**Sambrook J., Fritsch E.F. and Maniatis T. (1989):** *Molecular Cloning: a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

**Sanchez J.J. and Endicott P. (2006):** Developing multiplexed SNP assays with special reference to degraded DNA templates. *Nature Protocols* **1**: 1370-1378.

**Sanetra M. and Crozier R.H. (2003):** Patterns of population subdivision and gene flow in the ant *Nothomyrmecia macrops* reflected in microsatellite and mitochondrial DNA markers. *Molecular Ecology* **12**: 2281-2295.

**Sanjuan A. and Comesana A.S. (2002):** Molecular identification of nine commercial flaffish species by polymerase chain reaction – restriction fragment length polymorphism analysis of a segment of the cytochrome b region. *J Food Prot.* **65**(6): 1016-1023.

**San Mauro D., Gower D.J., Zardoya R. and Wilkinson M. (2006):** A hotspot of gene order rearrangement by tandem duplication and random loss in the vertebrate mitochondrial genome. *Mol. Biol. Evol.* **23**(1): 227-234.

**Sassa C., Kawaguchi K., Hirota Y. and Ishida M. (2004):** Distribution patterns of larval myctophid fish assemblages in the subtropical-tropical waters of the western North Pacific. *Fisheries Oceanography* **13**(4): 267-282.

**Sato Y., Sugie R., Tsuchiya B., Kameya T., Natori M. and Mukai K. (2001):** Comparison of the DNA extraction methods for polymerase chain rection amplification from formalin-fixed and paraffin-embedded tissues. *Diagnostic Molecular Pathology* **10**(4): 265-271.

**Saunders G.W. (2005):** Applying DNA barcoding to red macroalgae: a preliminary appraisal holds promise for future application. *Phil. Trans. R. Soc. B* **360** (1462): 1879-1888.

**Savioz A., Blouin J.-L., Guidi S., Antonarakis S.E. and Bouras C. (1997):** A method for the extraction of genomic DNA from human brain tissue fixed and stored in formalin for many years. *Acta Neuropathol* **93**: 408-413.

**Savolainen V., Cuenoud P., Spichiger R., Martinez M.D.P., Crevecoeur M. and Manen J-F. (1995):** The use of herbarium specimens in DNA phylogenetics: Evaluation and improvement. *Plant Systematics and Evolution* **197**: 87-98.

**Schander C. and Halanych K.M. (2003):** DNA, PCR and formalinized animal tissue – a short review and protocols. *Organisms diversity and evolution* **3**: 195-205.

**Schander C. and Willassen E. (2005):** What can biological barcoding do for marine biology? *Marine Biology Research* **1**: 79-83.

**Scheef E.A., Casler M.D. and Jung G. (2003):** Development of species-specific SCAR markers in bentgrass. *Crop Sci.* **43**: 345-349.

**Schiermeier Q. (2006):** A sea change. *Nature* **439**: 256-260.

**Schierwater B. and Ender A. (1993):** Diferent thermostable DNA polymerases may amplify different RAPD products. *Nucleic Acids Res.* **21**: 4647-4648.

**Schill R.O. (2007):** Comparison of different protocols for DNA preparation and PCR amplification of mitochondrial genes of tardigrades. *Journal of Limnology* **66**: 164-170.

**Schizas N.V., Street G.T., Coull B.C., Chandler G.T. and Quattro J.M. (1997):** An efficient DNA extraction method for small metazoans. *Molecular Marine Biology and Biotechnology* **6** (4): 381-383.

**Schlotterer C. (2000):** Evolutionary dynamics of microsatellite DNA. *Chromosoma* **109**: 365-371.

**Schlotterer C. (2004):** The evolution of molecular markers – just a matter fashion? *Nature Reviews/Genetics* **5**: 63-69.

**Schlotterer C. and Pemberton J. (1994):** The use of microsatellite for genetic analysis of natural populations. pp. 203-214 *in* Schierwater B., Streit B., Wagner G.P. and DeSalle R., editors. ***Molecular ecology and evolution: Approaches and applications***. Birkhauser Verlag, Basel, Boston, Berlin.

**Schwartz M. and Vissing J. (2002):** Paternal inheritance of mitochondrial DNA. *N Engl J Med* **347**: 576-580.

**Scoles D.R., Collette B.B. and Graves J.E. (1998):** Global phylogeography of mackerels of the genus *Scomber. Fishery Bulletin* **96**(4): 823-842.

**Scott M.P., Haymes K.M. and Williams S.M. (1992):** Parentage analysis using RAPD PCR. *Nucleic Acids Res.* **20**: 5493.

**Sefc K.M., Payne R.B. and Sorenson M.D. (2003):** Microsatellite amplification from museum feather samples: effects of fragment size and template concentration on genotyping errors. *The Auk* **120**(4): 982-989.

**Segarra-Moragues J.G. and Catalan P. (2003):** Life history variation between species of the relictual genus *Borderea* (Dioscoreaceae): phylogeography, genetic diversity, and population genetic structure assessed by RAPD markers. *Biological Journal of the Linnean Society* **80**: 483-498.

**Selkoe K.A. and Toonen R.J. (2006):** Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* **9**(5): 615-629.

**Sepp R., Szabo I., Uda H. and Sakamoto H. (1994):** Rapid techniques for DNA extraction from routinely processed archival tissue for use in PCR. *J Clin Pathol* **47**: 318-323.

**Sereno P.C. (2005):** The logical basis of phylogenetic taxonomy. *Syst. Biol.* **54**(4): 595-619.

**Serre D., Hofreiter M. and Paabo S. (2004):** Mutations induced by ancient DNA extracts? *Mol. Biol. Evol.* **21**(8): 1463-1467.

**Seyoum S., Denison S.H. and Tringali M.D. (2007):** Isolation and characterization of 13 polymorphic microsatellite loci for the Florida pompano, *Trachinotus carolinus*. *Molecular Ecology Notes* **7**(1): 141-143.

**Shaffer H.B., Fisher R.N. and Davidson C. (1998):** The role of natural history collections in documenting species declines. *TREE* **13**(1): 27-30.

**Sharma R., Singh M. and Sharma A. (2002):** Polymerase chain reaction: an emerging tool for research pharmacology. *Indian Journal of Pharmacology* **34**: 229-236.

**Shaw P.W., Arkhipkin A.I. and Al-Khairulla H. (2004):** Genetic structuring of Patagonian toothfish populations in the Southwest Atlantic Ocean: the effect of the Antarctic Polar Front and deep-water troughs as barriers to genetic exchange. *Molecular Ecology* **13**: 3293-3303.

**Shedlock A.M., Parker J.D., Cripsin D.A., Pietsch T.W. and Burmer G.C. (1992):** Evolution of the salmonid mitochondrial central region. *Molec Phylogenetics Evolut* **1**: 179-192.

**Shedlock A.M., Haygood M.G., Pietsch T.W. and Bentzen P. (1997):** Enhanced DNA extraction and PCR amplification of mitochondrial genes from formalin-fixed museum specimens. *BioTechniques* **22**(3): 394-400.

**Shi S-R., Cote R.J., Wu L., Liu C., Datar R., Shi Y., Liu D., Lim H. and Taylor C.R. (2002):** DNA extraction from archival formalin-fixed, paraffin-embedded tissue sections based on the antigen retrieval principle: heating under the influence of pH. *The Journal of Histochemistry and Cytochemistry* **50**(8): 1005-1011.

**Shi S-R., Datar R., Liu C., Wu L., Zhang Z., Cote R.J. and Taylor C.R. (2004):** DNA extraction from archival formalin-fixed, paraffin-embedded tissues: heat-induced retrieval in alkaline solution. *Histochem Cell Biol* **122**: 211-218.

**Shibata D. (1994):** Extraction of DNA from paraffin-embedded tissue for analysis by polymerase chain reaction: new tricks from an old friend. *Hum Pathol.* **25**(6): 561-563.

**Shields P.A. and Carlson S.R. (1996):** Effects of formalin and alcohol preservation on lengths and weights of juvenile sockeye salmon. *Alaska Fishery Research Bulletin* **3**(2): 81-93.

**Shikano T. and Taniguchi N. (2003):** DNA markers for estimation of inbreeding depression and heterosis in the guppy *Poecilia reticulata*. *Aquaculture Research* **34**(11): 905-911.

**Shiozawa D.K., Kudo J., Evans R.P., Woodward S.R. and Williams R.N. (1992):** DNA extraction from preserved trout tissues. *Great Basin Naturalist* **52**(1): 29-34.

**Shitara H., Hayashi J-I., Takahama S., Kaneda H. and Yonekawa H. (1998):** Maternal inheritance of mouse mtDNA followed by the prevention of subsequent paternal leakage. *Genetics* **148**: 851-857.

**Shitara H., Kaneda H., Sato A., Inoue K., Ogura A., Yonekawa H. and Hayashi J-I. (2000):** Selective and continuous elimination of mitochondria microinjected into mouse eggs from spermatids, but not from liver cells, occurs throught embryogenesis. *Genetics* **156**: 1277-1284.

**Sica M., Aceto S., Genovese A. and Gaudio L. (2002):** Analysis of five ancient equine skeletons by mitochondrial DNA sequencing. *Ancient Biomolecules* **4**(4): 179-184.

**Siebenaller J.F. (2003):** Pressure effects on the GTPase activity of brain membrane G proteins of deep-living marine fishes. *Comparative Biochemistry and Physiology Part B* **135**: 697-705.

**Siebenaller J.F. and Somero G.N. (1989):** Biochemical adaptation to the deep-sea. *CRC Critical Reviews in Aquatic Sciences* **1**: 1-25.

**Sikora J., Srbova A., Koukolik F. and Matej R. (2006):** Retrospective sequence analysis of the human PRNP gene from the formaldehyde-fixed paraffin-embedded tissues: report of two cases of Creutzfeldt-Jakob disease. *Folia Microbiol.* **51**(6): 619-625.

**Silberman J.D. and Walsh P.J. (1992):** Species identification of spiny lobster phyllosome larvae via ribosomal DNA analysis. *Mol. mar. Biol. Biotech.* **1**: 195-205.

**Simmons J.E. (1991):** Conservation problems of fluid-preserved collections. pp: 69-89 *in*: Cato P.S. and Jones C. (eds), ***Natural History Museums: Directions for Growth.*** Texas Tech University Press, Lubbock.

**Simon C., Frati F., Beckenbach A., Crespi B., Liu H. and Flook P. (1994):** Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Ann. Entomol. Soc. Am.* **87**: 651-701.

**Simon C., Buckley T.R., Frati F., Stewart J.B. and Beckenbach A.T. (2006):** Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase reaction primers for animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics* **37**: 545-579.

**Simon M.C., Gray D.I. and Cook N. (1996):** DNA extraction and PCR methods for the detection of *Listeria monocytogenes* in cold-smoked salmon. *Applied and Environmental Microbiology* **62**(3): 822-824.

**Singer M.F. (1982):** SINEs and LINEs: highly repeated short and long interspersed sequences in mammalian genomes. *Cell* **28**: 433-434.

**Sites J.W., Jr., and Marshall J.C. (2004):** Operational criteria for delimiting species. *Annu. Rev. Ecol. Evol. Syst.* **35**: 199-227.

**Siwoski A., Ishkanian A., Garnis C., Zhang L., Rosin M. and Lam W.L. (2002):** An efficient method for the assessment of DNA quality of archival microdissected specimens. *Mod Pathol* **15**(8): 889-892.

**Skage M. and Schander C. (2007):** DNA from formalin-fixed tissue: extraction or repair? That is question. *Marine Biology Research* **3**: 289-295.

**Skotarczak B., Przybos E., Wodecka B. and Maciejewska A. (2004):** Sibling species within *Paramecium jenningsi* revealed by RAPD. *Acta Protozoologica* **43**: 29-35.

**Skroch P. and Nienhuis J. (1995):** Impact of scoring error and reproducibility RAPD data on RAPD based estimates of genetic distance. *Theoretical and Applied Genetics* **91**: 1432-2242.

**Smit A.F.A. (1999):** Interspersed repeats and other mementos of transposable elements in mammalian genomes. *Current Opinion in Genetics and Development* **9**: 657-663.

**Smith C.A., Baeten J. and Taylor J.S. (1998):** The ability of a variety of polymerases to synthesize past site-specific cis-syn, trans-syn-II, (6-4), and Dewar photoproducts of thymidylyl-(3'--> 5')-thymidine. *J Biol Chem* **273**: 21933-21940.

**Smith J.S.C. and Williams J.G.K. (1994):** Arbitrary primer mediated fingerprinting in plants: Case studies in plant breeding, taxonomy and phylogeny. pp 5-15 *in* Schierwater B., Streit B., Wagner G.P. and DeSalle R., editors. ***Molecular ecology and evolution: Approaches and applications***. Birkhauser Verlag, Basel, Boston, Berlin.

**Smith M.A., Fisher B.L. and Hebert P.D.N. (2005):** DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: the ants of Madagascar. *Phil. Trans. R. Soc. B* **360**: 1825-1834.

**Smith T. and Waterman M. (1981):** Identification of common molecular subsequences. *J. Mol. Biol.* **147**: 195-197.

**Snel B., Bork P. and Huynen M.A. (1999):** Genome phylogeny based on gene content. *Nature Genetics* **21**: 108-110.

**Snelgrove P.V.R. and Smith C.R. (2002):** A riot of species in an environmental calm: the paradox of the species-rich deep sea. *Oceanography and Marine Biology Annual Review* **40**: 311–342.

**Snyder M., Fraser A.R., LaRoche J., Gartner-Kepkay K.E. and Zouros E. (1987):** Atypical mitochondrial DNA from the deep-sea scallop *Placopecten magellanicus*. *Proc. Natl. Acad. Sci. USA* **84**: 7595-7599.

**Sole M., Porte C. and Albaiges J. (2001):** Hydrocarbons, PCBs and DDT in the NW Mediterranean deep-sea fish *Mora moro*. *Deep-Sea Research I* **48**: 495-513.

**Solomon M.J. and Varshavsky A. (1985):** Formaldehyde-mediated DNA-protein crosslinking: A probe for *in vivo* chromatin structures. *Proc. Natl. Acad. Sci. USA* **82**: 6470-6474.

**Somero G.N. (1982):** Physiological and biochemical adaptations of deep-sea fishes: adaptive responses to the physical and biological characteristics of the abyss. pp. 257-278 *in* Ernst W.G. and Morin J.G., editors. ***The environment of the deep-sea. Rubey.*** **vol. II.** Prentice-Hall, Englewood Cliffs.

**Sommerfeldt A.D. and Bishop J.D.D. (1999):** Random amplified polymorphic DNA (RAPD) analysis reveals extensive natural chimerism in a marine protochordate. *Molecular Ecology* **8:** 885-890.

**Sorenson M.D., Cooper A., Paxinos E.E., Quinn T.W., James H.F., Olson S.L. and Fleischer R.C. (1999a):** Relationships of the extinct moa-nalos, flightless Hawaiian waterfowl, based on ancient DNA. *Proc. R. Soc. Lond. B* **266**: 2187-2193.

**Sorenson M.D., Ast J.C., Dimcheff D.E., Yuri T. and Mindell D.P. (1999b):** Primers for a PCR-based approach to mitochondrial genome sequencing in birds and other vertebrates. *Molecular Phylogenetics and Evolution* **12**(2): 105-114.

**Speers-Roesch B., Robinson J.W. and Ballantyne J.S. (2006):** Metabolic organisation of the spotted ratfish, *Hydrolagus colliei* (Holocephali: Chimaeriformes): Insight into the evolution of energy metabolism in the chondrichthyan fishes. *Journal of Experimental Zoology* **305A**: 631-644.

**Srinivasan M., Sedmak D. and Jewell S. (2002):** Effect of fixatives and tissue processing on the content and integrity of nucleic acids. *American Journal of Pathology* **161**(6): 1961-1971.

**Stanta G., Mucelli S.P., Petrera F., Bonin S. and Bussolati G. (2006):** A novel fixative improves opportunities of nucleic acids and proteomic analysis in human archive's tissues. *Diagn Mol Pathol.* **15**(2): 115-123.

**Steedman H.F. (1976):** ***Zooplankton fixation and preservation.*** UNESCO Press, Paris.

**Stephan W. and Cho S. (1994):** Possible role of natural selection in the formation of tandem-repetitive noncoding DNA. *Genetics* **136**: 333-341.

**Stephenson A.B. and Riley J.L. (1995):** Fixation and preservation of museum marine collections using formaldehyde/glutaraldehyde mixes. *Collection Forum* **11**(2): 58-68.

**Stiller M., Green R.E., Ronan M., Simons J.F., Du L., He W., Egholm M., Rothberg J.M., Keates S.G., Ovodov N.D., Antipina E.E., Baryshnikov G.F., Kuzmin Y.V., Vasilevski A.A., Wuenschell G.E., Termini J., Hofreiter M., Jaenicke-Despres V. and Paabo S. (2006):** Patterns of nucleotide misincorporations during enzymatic amplification and direct large-scale sequencing of ancient DNA. *PNAS* **103**(37): 13578-13584.

**Stockley B., Menezes G., Pinho M.R. and Rogers A.D. (2005):** Genetic populations structure in the black-spot sea bream (*Pagellus bogaraveo* Brunnich, 1768) from the NE Atlantic. *Marine Biology* **146**: 793-804.

**Stuart B.L., Dugan K.A., Allard M.W. and Kearney M. (2006):** Extraction of nuclear DNA from bone of skeletonized and fluid-preserved museum specimens. *Systematics and Biodiversity* **4**(2): 133-136.

**Su B., Wang Y-X., Lan H., Wang W. and Zhang Y. (1999):** Phylogenetic study of complete cytochrome *b* genes in musk deer (genus *Moschus*) using museum samples. *Molecular Phylogenetics and Evolution* **12**(3): 241-249.

**Suarez A.V. and Tsutsui N.D. (2004):** The value of museum collections for research and society. *BioScience* **54**(1): 66-74.

**Sugawara T., Terai Y. and Okada N. (2002):** Natural selection of the rhodopsin gene during the adaptive adiation of East African Great Lakes cichlid fishes. *Mol. Biol. Evol.* **19**(10): 1807-1811.

**Sunnucks P. (2000):** Efficient genetic markers for population biology. *TREE* **15**(5): 199-203.

**Swango K.L., Timken M.D., Chong M.D. and Buoncristiani M.R. (2006):** A quantitative PCR assay for the assessment of DNA degradation in forensic samples. *Forensic Science International* **158**: 14-26.

**Taberlet P. and Bouvet J. (1994):** Mitochondrial DNA polymorphism, phylogeography, and conservation genetics of the brown bear *Ursus arctos* in Europe. *Proc. R. Soc. Lond. B* **255**: 195-200.

**Taberlet P., Waits L.P. and Luikart G. (1999):** Noninvasive genetic sampling: look before you leap. *TREE* **14**(8): 323-327.

**Tang E.P.Y. (2006):** ***Path to effective recovering of DNA from formalin-fixed biological samples in natural history collections*: *Workshop Summary*.** The National Academies Press, Washington D.C. (http://www.nap.edu/catalog/11712.html).

**Tautz D. (1989):** Hypervariabile of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Res.* **17**: 6463-6471.

**Tautz D., Arctander A., Minelli A., Thomas R.H. and Vogler A.P. (2003):** A plea for DNA taxonomy. *Trends in Ecology and Evolution* **18**(2): 70-74.

**Taylor P.L. (1993):** ***GeneJockeyII – Sequence Processor.*** **– Operational Manual.** Biosoft, U.K.

**Taylor W.R. (1977):** Observation on specimen fixation. *Proceedings of the Biological Society of Washington* **90**: 753-763.

**Taylor W.R. (1981):** On preservation of color and collor patterns. *Curation Newsletter* **3**: 2-10.

**Tayutivutikul J., Pongprasert W., Royce L.A. and Ruangrit K. (2003):** Comparison of preservation techniques for silkworm (*Bombyx mori* L.) DNA based on polymerase chain reaction (PCR) products. *CMU Journal* **2**(2): 107-114.

**Telenius H., Carter N.P., Bebb C.E., Nordenskjold M., Ponder B.A. and Tunnacliffe A. (1992):** Degenerate oligonucleotide-primed PCR: General amplification of target DNA by a single degenerate primer. *Genomics* **13**(3): 718-725.

**Thalmann O., Hebler J., Poinar H.N., Paabo S. and Vigilant L. (2004):** Unreliable mtDNA data due to nuclear insertions: a cautionary tale from analysis of humans and other great apes. *Molecular Ecology* **13**: 321-335.

**Theodorakis C.W., Lee K-L., Adams S.M. and Law C.B. (2006):** Evidence of altered gene flow, mutation rate, and genetic diversity in redbreast sunfish from a pulp-mill-contaminated river. *Environ. Sci. Technol.* **40**: 377-386.

**Thibault-Botha D. and Bowen T. (2004):** Impact of formalin preservation on *Pleurobrachia bachei* (Ctenophora). *Journal of Experimental Marine Biology and Ecology* **303**(1): 11-17.

**Thiel H. (2003):** Anthropogenic impacts on the deep sea. pp. 427-471 *in* Tyler P.A. (ed.): *Ecosystems of the deep oceans* (*Ecosystems of the World* **28**). First edition. Elsevier Science B.V.

**Thistle D. (2003):** The deep-sea floor: an overview. pp. 5-37 471 *in* Tyler P.A. (ed.): *Ecosystems of the deep oceans* (*Ecosystems of the World* **28**). First edition. Elsevier Science B.V.

**Thomas R.H. (1994):** Molecules, museum and vouchers. *TREE* **9**(11): 413-414.

**Thomas W.K., Paabo S., Villablanca F.X. and Wilson A.C. (1990):** Spatial and temporal continuity of kangaroo rat population shown by sequencing mitochondrial DNA from museum specimens. *J. Mol. Evol.* **31**: 101-112.

**Thomas W.K., Vida J.T., Frisse L.M., Mundo M. and Baldwin J.G. (1997):** DNA sequences from formalin-fixed nematodes: integrating molecular and morphological approaches to taxonomy. *Journal of Nematology* **29**(3): 250-254.

**Thompson E.R., Herbert S.C., Forrest S.A. and Campbell I.G. (2005):** Whole genome SNP arrays using DNA derived from formalin-fixed, paraffin-embedded ovarian tumor tissue. *Human Mutation* **26**(4): 384-389.

**Threadgold J. and Brown T.A. (2003):** Degradation of DNA in artificially charred wheat seeds. *Journal of Archaeological Science* **30**: 1067-1076.

**Thrush S.F. and Dayton P.K. (2002):** Disturbance to marine benthic habitats by trawling and dredging: Implications for marine biodiversity. *Annu. Rev. Ecol. Syst.* **33**: 449-473.

**Thyagarajan B., Padua R.A. and Campbell C. (1996):** Mammalian mitochondria possess homologous DNA recombination activity. *J Biol Chem* **271**: 27536-27543.

**Tie J., Serizawa Y., Oshida S., Usami R. and Yoshida Y. (2005):** Individual identification by DNA polymorphism using formalin-fixed placenta with whole genome amplification. *Pathology International* **55**: 343-347.

**Tompa P. (2003):** Intrinsically unstructured proteins evolve by repeat expansion. *BioEssays* **25**: 847-855.

**Torres J.J., Belman B.W. and Childress J.J. (1979):** Oxygen consumption rates of midwater fishes as a function of depth of occurrence. *Deep-Sea Research* **26A**: 185-197.

**Toth J. and Biggin M.D. (2000):** The specificity of protein-DNA crosslinking by formaldehyde: *in vitro* and in *Drosophila* embryos. *Nucleic Acids Research* **28**(2): e4.

**Townson H., Harbach R.E. and Callan T.A. (1999):** DNA identification of museum specimens of the *Anopheles gambiae* complex: an evaluation of PCR as a tool for resolving the formal taxonomy of sibling species complexes. *Systematic Entomology* **24**: 95-100.

**Treasurer J.W. (1992):** Length and weight changes in 0+ perch, *Perca fluviatilis* L., following fixation in formalin. *Journal of Fish Biology* **41**: 1033-1036.

**Treberg J.R., Martin R.A. and Driedzic W.R. (2003):** Muscule enzyme activities in a deep-sea squaloid shark, *Centroscyllium fabricii*, compared with its shallow-living relative, *Squalus acanthias*. *Journal of Experimental Zoology* **300A**: 133-139.

**Tsaousis A.D., Martin D.P., Ladoukakis E.D., Posada D. and Zouros E. (2005):** Widespread recombination in published animal mtDNA sequences. *Mol. Biol. Evol.* **22**(4): 925-933.

**Tucker J.W., Jr. and Chester A.J. (1984):** Effects of salinity, formalin concentration and buffer on quality of preservation of southern flounder (*Paralichthys lethostigma*) larvae. *Copeia* **No. 2**: 981-988.

**Tuross N. (1994):** The biochemistry of ancient DNA in bone. *Experientia* **50**: 530-535.

**Tyler P.A. (1995):** Conditions for the existence of life at the deep-sea floor: An update. *Oceonography and Marine Biology Annual Review* **53**: 221-244.

**Tyler P.A. (2003):** Introduction. pp. 1-3 *in* Tyler P.A. (ed.): ***Ecosystems of the deep oceans*** (*Ecosystems of the World* **28**). First edition. Elsevier Science B.V.

**Umetsu K., Iwabuchi N., Yuasa I., Saitou N., Clark P.F., Boxshall G., Osawa M. and Igarashi K. (2002):** Complete mitochondrial DNA sequence of a tadpole shrimp (*Triops cancriformis*) and analysis of museum samples. *Electrophoresis* **23**: 4080-4084.

**Vachot A-M. and Monnerot M. (1996):** Extraction, amplification and sequencing of DNA from formaldehyde-fixed specimens. *Ancient Biomolecules* **1** (1): 3-16.

**Valiere N. and Taberlet P. (2000):** Urine collected in the field as a source of DNA for species and individual identification. *Molecular Ecology* **9**: 2150-2152.

**van Bortel W., Trung H.D., Roelants P., Harbach R.E., Backeljau T. and Coosemans M. (2000):** Molecular identification of *Anopheles minimus s.l.* beyond distinguishing the members of the species complex. *Insect Molecular Biology* **9**(3): 335-340.

**van Dam A.J. (2003):** DMDM-hydantoin: the promising result of a search for an alternative in fluid preservation of biological specimens. *Collection Forum* **18**: 104-115.

**van de Zande L. and Bijlsma R. (1995):** Limitations of the RAPD technique in phylogeny reconstruction in *Drosophila*. *J. Evol. Biol.* **8**: 645-656.

**van der Kuyl A.C., Kuiken C.L., Dekker J.T., Perizonius W.R.K. and Goudsmit J. (1995):** Nuclear counterparts of the cytoplasmatic mitochondrial 12S rRNA gene: a problem of ancient DNA and molecular phylogenies. *J. molec Evolut.* **40**: 652-657.

**Valasek M.A. and Repa J.J. (2005):** The power of real-time PCR. *Advances in Physiology Education* **29**: 151-159.

**Vass A.A. (2001):** Beyond the grave – understanding human decomposition. *Microbiology Today* **28**: 190-192.

**Vaughn K.C., DeBonte L.R. and Wilson K.G. (1980):** Organelle alternation as a mechanism for maternal inheritance. *Science* **208**: 196-198.

**Vilatersana R., Garnatje T., Susanna A. and Garcia-Jacas N. (2005):** Taxonomic problems in *Carthamus* (Asteraceae): RAPD markers and sectional classification. *Botanical Journal of the Linnean Society* **147**: 375-383.

**Vince A., Poljak M. and Seme K. (1998):** DNA extraction from archival Giemsa-stained bone-marrow slides: comparison of six rapid methods. *British Journal of Haematology* **101**: 349-351.

**Vink C.J., Thomas S.M., Paquin P., Hayashi C.Y. and Hedin M. (2005):** The effects of preservatives and temperatures on arachnid DNA. *Invertebrate Systematics* **19**: 99-104.

**Vinogradova N.G. (1997):** Zoogeography of the abyssal and hadal zones. *Advances in Marine Biology* **32**: 325-387.

**Volff J-N. (2005):** Genome evolution and biodiversity in teleost fish. *Heredity* **94**: 280-294.

**Volland W. (2004):** Formaldehyde. (http://scidid.bcc.ctc.edu/wv/form-0.htm).

**von Wurmb-Schwark N., Schwark T., Harbeck M. and Oehmichen M. (2004):** A simple duplex-PCR to evaluate the DNA quality of anthropological and forensic samples prior short tandem repeat typing. *Legal Medicine* **6**: 80-88.

**Vu N.T., Chaturvedi A.K. and Canfield D.V. (1999):** Genotyping for DQA1 and PM loci in urine using PCR-based amplification: Effects of sample volume, storage temperature, preservatives, and aging on DNA extraction and typing. *Forensic Science International* **102**: 23-34.

**Walden K.K.O. and Robertson H.M. (1997):** Ancient DNA from amber fossil bees? *Mol. Biol. Evol.* **14**(10): 1075-1077.

**Walsh P.S., Metzger D.A. and Higuchi R. (1991):** Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *BioTechniques* **10**: 506-513.

**Wan Q-H. and Fang S-G. (2003):** Species-specific "fingerprints" of deer in China. *Biodiversity and Conservation* **12**(6): 1253-1260.

**Wan Q-H., Wu H., Fujihara T. and Fang S-G. (2004):** Which genetic marker for which conservation genetic issue? *Electrophoresis* **25**(14): 2165-2176.

**Wan Q-H., Yu J-Q., Yong Y-G. and Fang S-G. (2006):** Reliable genetic analyses based on formalin-fixed materials: a new noninvasive sampling perspective for the giant panda. *Conservation Genetics* **7**: 467-471.

**Wandeler P., Hoeck P.E.A. and Keller L.F. (2007):** Back to the future: museum specimens in population genetics. *TRENDS in Ecology and Evolution* **22**(12): 634-642.

**Wang C., Arneson N.C.R., McCready M., McCready D. and Done S.J. (2005):** Investigation of molecular alternations in sentinel lymph node metastasis and primary breast cancer using array comparative genomic hybridization. *Proc Amer Assoc Cancer Res* **46**: Tumor Biology 38: Markers of Tumor Progression 1, Abstract #5572.

**Wang G., Maher E., Brennan C., Chin L., Leo C., Kaur M., Zhu P., Rook M., Wolfe J.L. and Makrigiorgos G.M. (2004a):** DNA amplification method tolerant to sample degradation. *Genome Research* **14**: 2357-2366.

**Wang G., Brennan C., Rook M., Wolfe J.L., Leo C., Chin L., Pan H., Liu W-H., Brendan P. and Makrigiorgos G.M. (2004b):** Balanced-PCR amplification allows unbiased identification of genomic copy changes in minute cell and tissue samples. *Nucleic Acids Research* **32**(9): e76.

**Wang W., Kumar P., Schwarz M., Malone G., Haworth A. and Kumar S. (1994):** PCR amplification of 40 years-old paraffin-embedded tumor tissues: Comparison of four different DNA extraction and purification methods. *International Journal of Oncology* **5**(3): 453-457.

**Ward R.D., Woodwark M. and Skibinski D.O.F. (1994):** A comparison of genetic diversity levels in marine, freshwater and anadromus fish. *Journal of Fish Biology* **44**: 213-232.

**Ward R.D. and Grewe P.M. (1995):** Appraisal of molecular genetic techniques in fisheries. pp. 1-28 *in* Carvalho G.R. and Pitcher T.J., editors. ***Molecular genetics in fisheries***. First edition. Chapman & Hall.

**Ward R.D., Zemlak T.S., Innes B.H., Last P.R. and Hebert P.D.N. (2005):** DNA barcoding Australia`s fish species. *Phil. Trans. R. Soc. B* **360**: 1847-1857.

**Waren A. (1983):** An anatomical description of *Eulima bilineata* Alder with remarks on and a revision of *Pyramidelloides* Nevill (Mollusca, Prosobranchia, Eulimidae). *Zoologica Scripta* **12**(4): 273-294.

**Waterman M.S. (1995):** ***Introduction to computational biology***. Chapman and Hall, London.

**Watts P.C., Thompson D.J., Allen K.A. and Kemp S.J. (2007):** How useful is DNA extracted from the legs of archived insects for microsatellite-based population genetic analyses? *J Insect Conserv* **11**: 195-198.

**Waugh R., McLean K., Flavell A.J., Pearce S.R., Kumar A., Thomas B.B.T. and Powell W. (1997):** Genetic distribution of Bare-1-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (S-SAP). *Mol Gen Genet* **253**: 687-694.

**Wayne R.K. and Jenks S.M. (1991):** Mitochondrial DNA analysis supports extensive hybridization of the endangered red wolf (*Canis rufus*). *Nature* **351**: 565-568.

**Wayne R.K., Leonard J.A. and Cooper A. (1999):** Full of sound and fury: The recent history of ancient DNA. *Annu. Rev. Ecol. Syst.* **30**: 457-477.

**Webb S.C., Fidler A. and Renault T. (2007):** Primers for PCR-based detection of ostreid herpes virus-1 (OsHV-1): application in a survey of New Zealand molluscs. *Aquaculture* **272**: 126-139.

**Weiss K.M., Buchann A.V., Daniel C. and Stoneking M. (1994):** Optimizing utilization of DNA from rare or archival anthropological samples. *Human Biology* **66**(5): 789-804.

**Weissensteiner T. and Lanchbury J.S. (1996):** Strategy for controlling preferential amplification and avoiding false negatives in PCR typing. *BioTechniques* **21**(6): 1102-1108.

**Wheeler W.C. (1994):** Sources of ambiguity in nucleic acid sequence alignment. Pages 323-352 *in* Schierwater B., Streit B., Wagner G.P. and DeSalle R., editors. ***Molecular ecology and evolution: Approaches and applications***. Birkhauser Verlag, Basel, Boston, Berlin.

**Wheeler W.C. and Gladstein D.M. (1993):** Malign: A multiple Sequence Alignment Program. Program and Documentation, version 1.92. New York.

**White T.J. (1996):** The future of PCR technology: diversification of technologies and applications. Reviews: *Trends in Biotechnology* **14**(12): 478-483.

**Whittier C.A., Dhar A.K., Stem C., Goodall J. and Alcivar-Warren A. (1999):** Comparison of DNA extraction methods for PCR amplification of mitochondrial cytochrome c oxidase subunit II (COII) DNA from primate fecal samples. *Biotechnology Techniques* **13**: 771-779.

**Wiens J.J. and Penkrot T.A. (2002):** Delimiting species using DNA and morphological variation and discordant species limits in spiny lizards (*Sceloporus*). *Syst. Biol.* **51**(1): 69-91.

**Wilkinson G.S. and Chapman A.M. (1991):** Length and sequence variation in evening bat D-loop mtDNA. *Genetics* **128**: 607-617.

**Will K.W. and Rubinoff D. (2004):** Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. *Cladistics* **20**: 47-55.

**Will K.W., Mishler B.D. and Wheeler Q.D. (2005):** The perils of DNA barcoding and the need for integrative taxonomy. *Syst. Biol.* **54**(5): 844-851.

**Willerslev E., Hansen A.J. and Poinar H.N. (2004):** Isolation of nucleic acids and cultures from fossil ice and permafrost. *TRENDS in Ecology and Evolution* **19**(3): 141-147.

**Willerslev E. and Cooper A. (2005):** Ancient DNA. *Proc. R. Soc. B* **272**: 3-16.

**Williams C., Ponten F., Moberg C., Soderkvist P., Uhlen M., Ponten J., Sitbon G. and Lundeberg J. (1999):** A high frequency of sequence alterations is due to formalin fixation of archival specimens. *American Journal of Pathology* **155**(5): 1467-1471.

**Williams J.G.K., Kubelik A.R., Livak K.J., Rafalski J.A. and Tingey S.A. (1990):** DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acid Research* **18**: 6531-6535.

**Williams S.T. and Reid D.G. (2004):** Speciation and diversity on tropical rocky shores: a global phylogeny of snails of the genus *Echinolittorina*. *Evolution* **58**(10): 2227-2251.

**Wilson I.G. (1997):** Inhibition and facilitation of nucleic acid amplification. *Applied and Environmental Microbiology* **63**: 3741-3751.

**Wilson K., Thorndyke M., Nilsen F., Rogers A. and Martinez P. (2005):** Marine systems: moving into the genomics era. *Marine Ecology* **26**: 3-16.

**Wilson R.R. Jr. (2001):** A new species of *Nezumia* (Gadiformes: Macrouridae) from Fieberling Guyot, eastern North Pacific Ocean. *Rev Biol Trop* **49**(Suppl 1): 29-37.

**Wilson R.R. and Attia P. (2003):** Interrelationships of the subgenera of *Coryphaenoides* (Teleostei: Gadiformes: Macrouridae): synthesis of allozyme, peptide mapping, and DNA sequence data. *Mol Phylogenet Evol.* **27**(2): 343-347.

**Winker K. (2004):** Natural history museums in a postbiodiversity era. *BioScience* **54**(5): 455-459.

**Wirgin I., Maceda L., Stabile J. and Mesing C. (1997):** An evaluation of introgression of Atlantic coast striped bass mitochondrial DNA in a Gulf of Mexico population using formalin-preserved museum collections. *Molecular Ecology* **6**: 907-916.

**Wisely S.M., Maldonado J.E. and Fleischer R.C. (2004):** A technique for sampling ancient DNA that minimizes damage to museum specimens. *Conservation Genetics* **5**: 105-107.

**Woodhead M., Russell J., Squirrell J., Hollingsworth P.M., Mackenzie K., Gibby M. and Powells W. (2005):** Comparative analysis of population genetic structure in *Athyrium distentifolium* (Pteridophyta) using AFLPs and SSRs from anonymous and transcribed gene regions. *Molecular Ecology* **14**: 1681-1695.

**Wu J., Yu Z., Song X., Wang Y. and Song L. (2005):** Rapid analysis of *Lateolabrax japonicus* DNA damage stressed by anthraquinone and SDBS. *Chinese Journal of Applied and Environmental Biology* **11**(4): 467-470.

**Wu L., Patten N., Yamashiro C.T. and Chui B. (2002):** Extraction and amplification of DNA from formalin-fixed, paraffin-embedded tissue. *Applied Immunohistochemistry and Molecular Morphology* **10**: 269-272.

**Yakes F.M. and Van Houten B. (1997):** Mitochondrial DNA damage is more extensive and persists longer than nuclear DNA damage in human cells following oxidative stress. *Proc. Natl. Acad. Sci. USA* **94**: 514-519.

**Yamashita S. (2007):** Heat-induced antigen retrieval: Mechanisms and application to histochemistry. *Progress in Histochemistry and Cytochemistry* **41**: 141-200.

**Yamashita S. and Okada Y. (2005):** Mechanisms of heat-induced antigen retrieval: analyses *in vitro* employing SDS-PAGE and immunohistochemistry. *Journal of Histochemistry and Cytochemistry* **53**(1): 13-21.

**Yang D.Y. and Watt K. (2005):** Contamination controls when preparing archaeological remains for ancient DNA analysis. *Journal of Archaeological Science* **32**: 331-336.

**Yang H., Golenberg E.M. and Shoshani J. (1997a):** Proboscidean DNA from museum and fossil specimens: An assessment of ancient DNA extraction and amplification techniques. *Biochemical Genetics* **35**: 165-179.

**Yang H., Golenberg E.M. and Shoshani J. (1997b):** A blind testing design for authenticating ancient DNA sequences. *Molecular Phylogenetics and Evolution* **7**(2): 261-265.

**Yang Z-J., Liu C., Feng J., Li G-R., Zhou J-P., Deng K-J. and Ren Z-L. (2006):** Studies on genome relationship and species-specific PCR marker for *Dasypyrum breviaristatum* in Triticeae. *Hereditas* **143**: 47-54.

**Yoder M., Tandingan De Ley I., King I.W., Mundo-Ocampo M., Mann J., Blaxter M., Poiras L. and De Ley P. (2006):** DESS: a versatile solution for preserving morphology and extractable DNA of nematodes. *Nematology* **8**(3): 367-376.

**Yokoyama S. (2000):** Molecular evolution of vertebrate visual pigments. *Prog. Retin. Eye Res.* **19**: 385-419.

**Young C.M. (2003):** Reproduction, development and life-history traits. pp. 381-426 *In* Tyler P.A. (ed.): ***Ecosystems of the deep oceans*** (*Ecosystems of the World* **28**). First edition. Elsevier Science B.V.

**Yue G.H. and Orban L. (2001):** Rapid isolation of DNA from fresh and preserved fish scales for polymerase chain reaction. *Marine Biotechnology* **3**: 199-204.

**Yue G.H., Ong D., Wong C.C., Lim L.C. and Orban L. (2003):** A strain-specific and sex-associated STS marker for Asian arowana (*Scleropages formosus*, Osteoglossidae). *Aquaculture Research* **34**: 951-957.

**Zaragueta-Bagils R., Lavoue S., Tillier A., Bonillo C. and Lecointre G. (2002):** Asessment of otocephalan and protacanthopterygian concepts in the light of multiple molecular phylogenies. *C.R. Biologies* **325**: 1191-1207.

**Zardus J.D., Etter R.J., Chase M.R., Rex M.A. and Boyle E.E. (2006):** Bathymetric and geographic population structure in the pan-Atlantic deep-sea bivalve *Deminucula atacellana* (Schenck, 1939). *Molecular Ecology* **15**: 639-651.

**Zezina O.N. (1997):** Biogeoraphy of the bathyal zone. *Advances in Marine Biology* **32**: 389-426.

**Zhang D-X. and Hewitt G.M. (1996):** Nuclear integrations: challenges for mitochondrial DNA markers. *Trends Ecol Evolut* **11**: 247-251.

**Zhang L., Cui X., Schmitt K., Hubert R., Navidi W. and Arnheim N. (1992):** Whole genome amplification from a single cell: Implications for genetic analysis. *Proc. Natl. Acad. Sci. USA* **89**: 5847-5851.

**Zouros E., Freeman K.R., Ball A.O. and Pogson G.H. (1992):** Direct evidence for extensive paternal mitochondrial DNA inheritance in the marine mussel *Mytilus*. *Nature* **359**: 412-414.

**Zsikla V., Baumann M. and Cathomas G. (2004):** Effect of buffered formalin on amplification of DNA from paraffin wax embedded small biopsies using real-time PCR. *J Clin Pathol* **57**: 654-656.

**Zwanenburg K.C.T., Bentzen P. and Wright J.M. (1992):** Mitochondrial DNA differentiation in western North Atlantic population of haddock (*Melanogrammus aeglefinus*). *Can. J. Fish. Aquat. Sci.* **49**: 2527-2537.

APPENDIX 1,2,3 and 4 have not been copied on instruction from the university

THIS THESIS CONTAINS A CD WHICH WE ARE NOT PERMITTED TO COPY

PLEASE CONTACT THE UNIVERSITY IF YOU WISH TO SEE THIS MATERIAL